

Result No.	Score	%		DB	ID	Description
		Match	Length			
1	1344	100.0	2565	6	AX063459	AX063459 Sequence
C 2	129.6	9.6	7218	6	I6494	I6494 Sequence 14
C 3	104	7.7	82781	9	AC103691	AC103691 Homo sapi
C 4	103.4	7.7	188507	10	AC115738	AC115738 Mus muscu
5	103	7.7	10732	6	E32986	E32986 Gene encodi
6	102.2	7.6	265985	2	AC0878226	AC0878226 Mus muscu
7	101.8	7.6	267182	2	AC116720	AC116720 Mus muscu
C 8	100.8	7.5	189129	2	AL691432	AL691432 Homo sapi
9	100.2	7.5	110000	2	AC103465	Continuation (4 of
10	99.6	7.4	132080	9	AC069368	AC069368 Homo sapi
11	98.2	7.3	303091	2	AC084799	AC084799 Mus muscu
12	97.8	7.3	196204	9	AC073957	AC073957 Homo sapi
C 13	97.6	7.3	130540	2	AC079417	AC079417 Mus muscu
C 14	97.4	7.2	155840	2	AC006495	AC006495 Drosophil
C 15	97.2	7.2	177277	10	AC147269	AC147269 Mus muscu
C 16	97.2	7.2	182065	10	AC138529	AC138529 Mus muscu
C 17	96.8	7.2	216370	10	AC091332	AC091332 Mus muscu
C 18	96.6	7.2	216501	10	AC107795	AC107795 Mus muscu
19	96.4	7.2	208161	2	AC074145	AC074145 Mus muscu

[illegible]

Db 1516 GAGGCCCAACACCAAGCCCAAGAAACAGAAAGCCCGCGCAGCAAGACCAACAC 1515

QY 1321 AGCAGAGAAGAGAGAGACCCCGC 1344

Db 1576 AGCAGAGAAGAGAGAGACCCCGC 1599

RESULT 2

I66494 7218 bp DNA linear PAT 28-DEC-1997

LOCUS Sequence 14 from patent US 5670367.

ACCESSION I66494

VERSION I66494.1 GI:2724471

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..7218

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 9.6%; Score 129.6; DB 6; Length 7218;

Best Local Similarity 1.6%; Pred. No. 1.4e-17;

Matches 6; Conservative 289; Mismatches 83; Indels 0; Gaps 0;

QY 870 ACAGAGACAGAGAAACAGCAGGAGAGACCAGAAAAGGGACCGCAGGAGAAAGGCCAGC 929

Db 1435 ACER 1377

QY 930 AGAAGAGAAAGCAGCCGGAGGAAAGCCAGACAGCGGAAAGCAAGCGAAAGACAGCGGACG 989

Db 1376 RRRER 1317

QY 990 CGAACACAGACAAAGAACGCGCAGGACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049

Db 1316 RRRER 1257

QY 1050 CAGGGGGGGGAAAAGGAGCCGAAAGAACCCAGCCACCCGAGAGAGAGAGAGAGAGAGAG 1109

Db 1256 RRRER 1197

QY 1110 AAAAGGGGCAACAAAGGGGAGCGGCAAGCAACACACCCCAACGAAAGAGAGAGAGAGAG 1169

Db 1196 RRRER 1137

QY 1170 AAGCGCGCGAGAGAAAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1229

Db 1136 RRRER 1077

QY 1230 CAAAGGAGACCCCAAG 1247

Db 1076 RRRER 1059

RESULT 3

AC103691/c

LOCUS Homo sapiens chromosome 15, clone RP11-959H21, complete sequence.

DEFINITION

ACCESSION AC103691

VERSION AC103691.2 GI:20377033

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 82781)

Birren,B.,Linton,L., Nussbaum,C. and Lander,E.

Faró, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 188507)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faró, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 188507)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faró, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 23, 2003 this sequence version replaced gi:37806517.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22816
Center clone name: 474_P_24

FEATURES
source
Location/Qualifiers
1..188507
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP23-474P24"
/clone_lib="RPCI-23 Female Mouse BAC"
1..11149
/note="wgs end extension
clone_end:SP6"
complement(1..396)
/rpt_family="L1"
complement(430..1594)
/rpt_family="Lx9"
1649..1744
/rpt_family="(TA)n"
1905..2073
/rpt_family="L1"
2189..2847
/rpt_family="Lx"
complement(4661..4821)
/rpt_family="MIR"
4833..4980
/rpt_family="B1_MM"
complement(5077..5164)
/rpt_family="MIR"
5206..5227
/rpt_family="AT_rich"
7500..7535
/rpt_family="(ATGGTC)n"
8932..9028
/rpt_family="(TG)n"
complement(9055..9331)
/rpt_family="B4A"
9714..9743
/rpt_family="(TTTG)n"
complement(9857..10851)
/rpt_family="Lx9"
10858..10887
/rpt_family="(TC)n"
11150..11155
/note="clone boundary
clone_end:SP6
site:ECORI"
11274..11299
/rpt_family="(A)n"
12002..12056
/rpt_family="(TCTA)n"
complement(12805..12904)
/rpt_family="PB1D10"
13168..13203
/rpt_family="(TC)n"
13205..13239
/rpt_family="(TG)n"
13242..13261
/rpt_family="(GA)n"
15794..15865
/rpt_family="(TGGA)n"
16699..16747
/rpt_family="AT_rich"
complement(17093..17202)
/rpt_family="PB1D9"
complement(17224..17358)

```

/rpt_family="B1_MM"
17362..17395
/rpt_family="TTTG)n"
complement(17404..17491)
/rpt_family="B1_MM"
complement(17415..17516)
/rpt_family="B4A"
17589..17724
/rpt_family="B1_MM"
18097..18215
/rpt_family="(TA)n"
complement(18216..18300)
/rpt_family="LMC4"
18327..18392
/rpt_family="(TA)n"
18400..18430
/rpt_family="(TCTA)n"
19720..19742
/rpt_family="(A)n"
19839..19865
/rpt_family="(TTTA)n"
21970..22017
/rpt_family="AT_rich"
22266..22334
/rpt_family="CT-rich"
complement(22968..23114)
/rpt_family="B1_MM"
23936..24134
/rpt_family="MLTLA"
24437..24681
/rpt_family="B4A"
24715..24755
/rpt_family="(CA)n"
24755..24825
/rpt_family="(GA)n"

```

	Query Match	7.7%; Score 103.4; DB 10; Length 188507;			
	Best Local Similarity	45.5%; Pred. No. 5.4e-12;			
	Matches 446; Conservative	0; Mismatches 531; Indels 4; Gaps 2;			
OY	284	AGGAAAGCGACCAAAAGAAGCAACAGAAACGGCCGCCAGCAGCGAACCAACAGAGCAGAA	343		
Ddb	102499	AGAGAGAAAGAGAGAGAGAAAGAAAGAGAGAGAGAAAGAAAGAAAGAGAG	102440		
OY	344	GCCCCGAGCGCCCCGAGAGCGAGAGCGCACACACACCAGAGGGCCCCCGCGAAGCA	403		
Ddb	102439	AGAAAAGAAAGAAAGAAAGAGAGAGAAAGAGAGAGAGAGAGAAAGAGATAGAG	102380		
OY	404	GGAAACGAGAAACCGCGGCAGAGACACACCGGAAGGAGGACCCAGGAGAGAACAGAGAGAG	463		
Ddb	102379	GAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAAAGAAAGAAAGAGAGAGAGAGAG	102320		
OY	464	CACCGCGGAGAGAGAGAGAGCGGAGAGGAGCCAGCAGAGAGGCCCGCCAGCGCCCACA	523		
Ddb	102319	AAAGAGAGAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	102260		
OY	524	CCCAGCAGCCGGAAGAGAAACCAAGAGAGGCCACAGGCGAGCCCAAGGAAGCCGAGAGAG	583		
Ddb	102259	AGAAAGAAAGAAAGAAAG	102200		
OY	584	--GGCCGCGGAGGAGAGAAACCGCGCAACACCGGCGAGAGACAAAAAGCAGAGCCCAAAG	641		
Ddb	102199	BAAAGAAAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	102140		
OY	642	AAGCAGCAAAAGACCGAAGACCGCCAGAGCCCAAGCCCAACCAAGAAACACACGGA	701		
Ddb	102139	AGAAAGAAAGAA	102080		
OY	702	CGAGAGGNAACCCAACCAACGCGCCGAGACCGAGGCCCGCGGCGAGAGAGAACCAACCA	761		
Ddb	102079	AGAAAGAAAGAA	102020		
OY	762	CACCGCGGAGGCCAAAAAGAGACCAACCCCAACACAGAGGAGAGAGAGAACCAACGAGAA	821		

```

-----
Project Information
Center Project Name: 0
Center clone name: RPCI-23_103M13
-----
Summary Statistics
Consensus quality: 151438 bases at least Q40
Consensus quality: 173959 bases at least Q30
Consensus quality: 185196 bases at least Q20
Estimated insert size: 220300; agarose-fp estimation
Estimated insert size: 258985; sum-of-contigs estimation
Quality coverage: 3.77 in Q20 bases; agarose-fp estimation
Quality coverage: 3.21 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 72 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1095: contig of 1095 bp in length
* 1096 1195: gap of unknown length
* 1196 2389: contig of 1194 bp in length

```

* 2390	2489: gap of unknown length	* 47281	48753: contig of 1473 bp in length
* 2490	3686: contig of 1197 bp in length	* 48754	48853: gap of unknown length
* 3687	3786: gap of unknown length	* 48854	50036: contig of 1183 bp in length
* 3787	4936: contig of 1210 bp in length	* 50037	50136: gap of unknown length
* 4937	5096: gap of unknown length	* 50137	51270: contig of 1134 bp in length
* 5097	6206: contig of 1110 bp in length	* 51271	51370: gap of unknown length
* 6207	6306: gap of unknown length	* 51371	52534: contig of 1164 bp in length
* 6307	7431: contig of 1125 bp in length	* 52535	52634: gap of unknown length
* 7432	7531: gap of unknown length	* 52635	53751: contig of 1117 bp in length
* 7532	8614: contig of 1083 bp in length	* 53752	53851: gap of unknown length
* 8615	8714: gap of unknown length	* 53852	54995: contig of 1144 bp in length
* 8715	9848: contig of 1134 bp in length	* 54996	55095: gap of unknown length
* 9849	9948: gap of unknown length	* 56244	56243: contig of 1148 bp in length
* 11138	11138: contig of 1190 bp in length	* 56344	56343: gap of unknown length
* 11139	11238: gap of unknown length	* 56345	57370: contig of 1027 bp in length
* 12302	12302: contig of 1064 bp in length	* 57371	57470: gap of unknown length
* 12402	12402: gap of unknown length	* 57471	57470: gap of unknown length
* 12403	13537: contig of 1135 bp in length	* 58714	58713: contig of 1243 bp in length
* 13538	13637: gap of unknown length	* 58814	58813: gap of unknown length
* 13638	14822: contig of 1185 bp in length	* 59844	59843: contig of 1030 bp in length
* 14823	14922: gap of unknown length	* 59944	59943: gap of unknown length
* 14923	16057: contig of 1135 bp in length	* 61020	61019: contig of 1076 bp in length
* 16058	16157: gap of unknown length	* 61120	61119: gap of unknown length
* 16158	17331: contig of 1174 bp in length	* 62327	62326: contig of 1207 bp in length
* 17332	17431: gap of unknown length	* 62427	62426: gap of unknown length
* 17432	18574: contig of 1143 bp in length	* 63762	63761: contig of 1335 bp in length
* 18575	18674: gap of unknown length	* 63862	63861: gap of unknown length
* 18675	19805: contig of 1131 bp in length	* 64973	64972: contig of 1111 bp in length
* 19806	19905: gap of unknown length	* 65073	65072: gap of unknown length
* 19906	21031: contig of 1126 bp in length	* 66225	66224: contig of 1152 bp in length
* 21032	21131: gap of unknown length	* 66325	66324: gap of unknown length
* 21132	22252: contig of 1121 bp in length	* 67395	67394: contig of 1070 bp in length
* 22253	22352: gap of unknown length	* 67495	67494: gap of unknown length
* 22353	23500: contig of 1148 bp in length	* 68560	68559: contig of 1065 bp in length
* 23501	23600: gap of unknown length	* 68660	68659: gap of unknown length
* 23601	24737: contig of 1137 bp in length	* 70795	70794: contig of 2135 bp in length
* 24738	24837: gap of unknown length	* 70895	70894: gap of unknown length
* 24838	26009: contig of 1172 bp in length	* 73759	73758: contig of 2864 bp in length
* 26010	26109: gap of unknown length	* 73859	73858: gap of unknown length
* 26110	27231: contig of 1122 bp in length	* 75387	75386: contig of 1528 bp in length
* 27232	27331: gap of unknown length	* 75487	75486: gap of unknown length
* 27332	28487: contig of 1156 bp in length	* 78567	78566: contig of 3081 bp in length
* 28488	28587: gap of unknown length	* 78668	78667: gap of unknown length
* 28588	29752: contig of 1165 bp in length	* 83212	83211: contig of 4544 bp in length
* 29753	29852: gap of unknown length	* 83312	83311: gap of unknown length
* 29853	30944: contig of 1092 bp in length	* 88001	88000: contig of 4689 bp in length
* 30945	31044: gap of unknown length	* 88101	88100: gap of unknown length
* 31045	32225: contig of 1181 bp in length	* 94341	94340: contig of 6240 bp in length
* 32226	32325: gap of unknown length	* 94441	94440: gap of unknown length
* 32326	33466: contig of 1141 bp in length	* 105004	105003: contig of 10563 bp in length
* 33467	33566: gap of unknown length	* 105104	105103: gap of unknown length
* 33567	34645: contig of 1079 bp in length	* 112337	112336: contig of 7233 bp in length
* 34646	34745: gap of unknown length	* 112437	112436: gap of unknown length
* 34746	35814: contig of 1069 bp in length	* 126184	126183: contig of 13747 bp in length
* 35815	35914: gap of unknown length	* 126284	126283: gap of unknown length
* 35915	37065: contig of 1151 bp in length	* 136889	136888: contig of 10605 bp in length
* 37066	37165: gap of unknown length	* 136989	136988: gap of unknown length
* 37166	38395: contig of 1230 bp in length	* 149641	149641: contig of 12653 bp in length
* 38396	38495: gap of unknown length	* 149742	149741: gap of unknown length
* 38496	39843: contig of 1348 bp in length	* 168566	168565: contig of 18824 bp in length
* 39844	39943: gap of unknown length	* 168666	168665: gap of unknown length
* 39944	41035: contig of 1092 bp in length	* 184977	184976: contig of 16311 bp in length
* 41036	41135: gap of unknown length	* 185077	185076: gap of unknown length
* 41136	42239: contig of 1104 bp in length	* 203645	203645: contig of 18569 bp in length
* 42240	42339: gap of unknown length	* 203646	203646: gap of unknown length
* 42340	43436: contig of 1097 bp in length	* 229532	229531: contig of 25786 bp in length
* 43437	43536: gap of unknown length	* 229532	229531: gap of unknown length
* 43537	44696: contig of 1160 bp in length	* 229632	229631: contig of 36354 bp in length.
* 44697	44796: gap of unknown length		
* 44797	45930: contig of 1134 bp in length		
* 45931	46030: gap of unknown length		
* 46031	47180: contig of 1150 bp in length		
* 47181	47280: gap of unknown length		

FEATURES
source
1. .265985
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="16"

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 10466: contig of 10466 bp in length
* 10467: gap of 100 bp
* 10566: contig of 1064 bp in length
* 11630: gap of 100 bp
* 11631: contig of 1889 bp in length
* 11731: gap of 100 bp
* 13620: contig of 1328 bp in length
* 13720: gap of 100 bp
* 15047: contig of 1415 bp in length
* 15148: gap of 100 bp
* 16562: contig of 1415 bp in length
* 16563: gap of 100 bp
* 16663: contig of 2312 bp in length
* 18974: gap of 100 bp
* 19075: contig of 4288 bp in length
* 23362: gap of 100 bp
* 23463: contig of 1814 bp in length
* 25276: gap of 100 bp
* 25277: contig of 6908 bp in length
* 32384: gap of 100 bp
* 32385: contig of 6718 bp in length
* 32385: gap of 100 bp
* 39103: contig of 10274 bp in length
* 39203: gap of 100 bp
* 49476: contig of 8375 bp in length
* 49477: gap of 100 bp
* 49577: contig of 8375 bp in length
* 57952: gap of 100 bp
* 58051: contig of 8097 bp in length
* 66148: gap of 100 bp
* 66149: contig of 14948 bp in length
* 66249: gap of 100 bp
* 81196: contig of 10117 bp in length
* 81297: gap of 100 bp
* 91413: contig of 10117 bp in length
* 91513: gap of 100 bp
* 117047: contig of 25534 bp in length
* 117048: gap of 100 bp
* 117147: contig of 48681 bp in length
* 165828: contig of 100 bp
* 165928: gap of 100 bp
* 221922: contig of 55994 bp in length
* 222023: gap of 100 bp
* 227182: contig of 45160 bp in length.

FEATURES

source
1..267182
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_lib="RP23-223N19"
/clone="RP23-223N19" Female Mouse BAC"
1..10466
/note="assembly_fragment"
clone_end:SP6
vector_side:left
10567..11630
/note="assembly_fragment"
11731..13619
/note="assembly_fragment"
13720..15047
/note="assembly_fragment"
15148..16562
/note="assembly_fragment"
16663..18974
/note="assembly_fragment"
19075..23362
/note="assembly_fragment"
23463..25276
/note="assembly_fragment"
25377..32284
/note="assembly_fragment"
32385..39102
/note="assembly_fragment"
39203..49476

misc_feature
49577..57951
/note="assembly_fragment"
58052..56148
/note="assembly_fragment"
66249..81196
/note="assembly_fragment"
81297..91413
/note="assembly_fragment"
91514..117047
/note="assembly_fragment"
117148..165828
/note="assembly_fragment"
165929..221922
/note="assembly_fragment"
222023..267182
/note="assembly_fragment"
clone_end:17
vector_side:right"

ORIGIN

Query Match 7.6% Score 101.8; DB 2; Length 267182;
Best Local Similarity 43.0%; Pred. No: 1.2e-11;
Matches 443; Conservative 0; Mismatches 567; Indels 21; Gaps 4;
QY 284 AGGAAGCGACCCMAAAGAAAGCAACAGAAACGCGCCGCCACCGACCAACAGAGACGAGCA 343
Db 164908 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 164967
QY 344 G-CCCCGAGCGCCCGGAG 402
Db 164968 GAGGAAGAAAG 165027
QY 403 AGGAACAGAAACCCAGGGGAG 462
Db 165028 AGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 165087
QY 463 GCACCGCGGAGCAGAGA--GAGGACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520
Db 165088 AAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 165147
QY 521 ACACCCAGCAGCCCGGAG 580
Db 165148 AGGAAGAAAG 165207
QY 581 AAGGCGCGGAGGAG 640
Db 165208 AGAAGGAAAG 165267
QY 641 GAAAGCAGCAAAAGACCGGAG 690
Db 165268 GAAAGGAAAG 165327
QY 691 AAGACACCGGAG 750
Db 165328 AGAAGAAAG 165387
QY 751 GAGAAACACCCAGCCCGGAG 810
Db 165388 GAAGAGAGAAAG 165447
QY 811 AACACCCAGGAGCCAAAGGAG 870
Db 165448 AAGAGAGGAG 165507
QY 871 CAGAGACAG 930
Db 165508 AGAAGAAAG 165567
QY 931 GAAAGGAAAGGACCGCGGAG 990
Db 165568 GGAAGAGGAG 165627

[illegible]

Db 169183 CAGAGAGAAACAGAGACAGACAGAGAGAGCGACAGAGAGACAGACAGAGAGAGA 169124
 Qy 1201 GAAAAAAGGCCCGAGGGCGAGCCAGCCAAAGAGAAACCCCAA 1246
 Db 169123 GCAGAACAGGGAGAGACAGAGAGAGACACAGAGACAGAGAGCCAAA 169078

RESULT 9
 AC103465_3
 WPCOMMENT
 Sequence split into 5 fragments LOCUS AC103465 Accession AC103465

Fragment Name	Begin	End
AC103465_0	1	110000
AC103465_1	100001	210000
AC103465_2	200001	310000
AC103465_3	300001	410000
AC103465_4	400001	452732

Continuation (4 of 5) of AC103465 from base 300001 (AC103465 Rattus norvegicus cl

Query Match	7.5%;	Score 100.2;	DB 2;	Length 110000;
Best Local Similarity	45.7%;	Pred. No. 2.7e-11;		
Matches 482;	Conservative	0;	Mismatches 553;	Indels 20;
Gaps	3;			
Qy 204	ACAGGAGGCCCGGCAACCGAGGAGGAGGAGAAACCCCGGACACAGCCCGGGACAG	263		
Db 16612	AAAGCAGAAGCAGAAAAGCAGAAGCAGGAGAGAAAGAGCAGAAACAAACAGAGCAAAAG	16671		
Qy 264	CCGGAGCCCGCGGCAGCAGGAGCGGACCCGACCAAAAGAAAGAAACAGAAACCGGCCGACCA	323		
Db 16672	CAGAAAGCAAAAAGAGAGACGAGAAAAGAGAGAGTAGAACHAGCAGAGCAGAAAAGAGAA	16731		
Qy 324	CGGACCAACGAGAGCAGAAAGCCCGAGCGCCCGGAGAGCAGAGAGCGACACACACAC	383		
Db 16732	GCAGGAAGAAAGAGCGCAAAATCAGAACGAGAAAGCAGAGCAAGAAAAACAGAGCAG	16791		
Qy 384	CAGAGGCCCGCGCGAAGCAGGAAACGAGAAACCGAGGGCAGGACCAACCCGAGGAGACC	443		
Db 16792	AAAAGCAGAAGCAGATATAACAGAAAGCAGAAAAGCAGAAGCAGATACAGAGAGAGAA--	16849		
Qy 444	AGCCAGAGAAACAGGAGAGACCGCGGAGCAGAGAGAGCGAGAGGACCAACGAGCAGA	503		
Db 16850	AAGCAGAGCAGGAAAACAGAACGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	16909		
Qy 504	AGSCCCACGAGCGCCACACCCACGACCCCGGAGAGAAAGAAACCAAGAGAGCCACAGGCG	563		
Db 16910	GAAGCAGAGCAGAGCAGAA-----AAGGAGAGCAGAAAAGCAGAAAGCA 16955			
Qy 564	AGCCAGAGAAACCGAGAAAGCGCGGAGCAGAGAGAGCGAGAGGACCAACGAGCAGAC	623		
Db 16956	GGAGCAGAGCCGAAAAGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	17015		
Qy 624	AAAAAGCAGACCAAGAGAGCAGAAAGACCCGAGAGCCGAGAGCCCAAGAGCCCA	683		
Db 17016	AGAAAGTAGAGCCGAAAAGCAGAGCAGGAAAGAGAGAGCAGAAAGAGCAGAAAGCAGAA	17075		
Qy 684	ACCAAGAAAGACCGGACAGAGAGAAACCCAAACAGGCGCGAGACCGGAGCCACCGGG	743		
Db 17076	AAGCAGAAACAGAAAGAGAGCAGAAAGAGCAGAAAGAGAGAGAGAGAGAGAGAGAG	17135		
Qy 744	CCGAGAGGAGAAACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	803		
Db 17136	GCAGAGCAGAAAGAGCAG	17195		
Qy 804	GAAGGAGAAACACCGAG	863		
Db 17196	GAAAGAGCAGAAAG	17255		
Qy 864	CGCCAGCAG	923		
Db 17256	AAAGCAG	17311		
Qy 924	GCCAGCAGAAAG	983		

Submitted (28-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 132080)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chararo, B., Choepel, Y., Collangola, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

```

Location/Qualifiers
1. .132080
   /organism="Homo sapiens"
   /mol_type="Genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="15"
   /map="15"
   /clone="CTD-2017F17"
   /clone_lib="CITD1 Human BAC"
438. 602
   /rpt_family="MERSB"
615. 663
   /rpt_family="L3"
1348. 1492
   /rpt_family="FLAM C"
complement(1670..2021)
   /rpt_family="LINEB4"

```

[illegible]

```

Db      16821 GAAGAGAACTAGAGA 16838
RESULT 11
AC084799
LOCUS   AC084799
DEFINITION AC084799 303091 bp DNA linear HTG 17-NOV-2000
          Mus musculus chromosome 16 clone RP23-197M9, WORKING DRAFT
          SEQUENCE, 101 unordered pieces.
ACCESSION AC084799
VERSION   AC084799.1 GI:11192127
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 303091)
AUTHORS   DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2000) Production Sequencing Facility, DOE Joint
          Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT   -----Genome Center
          Center: Joint Genome Institute
          Center Code: JGI
          Web site: http://www.jgi.doe.gov
          -----
Project Information
Center Project Name: 0
Center clone name: RPCI-23_197M9
-----
Summary Statistics
Consensus quality: 152568 bases at least Q40
Consensus quality: 175579 bases at least Q30
Consensus quality: 186949 bases at least Q20
Estimated insert size: 198300; agarose-fp estimation
Estimated insert size: 293091; sum-of-contigs estimation
Quality coverage: 4.85 in Q20 bases; agarose-fp estimation
Quality coverage: 3.28 in Q20 bases; sum-of-contigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 101 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1117 1116: contig of 1116 bp in length
1217 1216: gap of unknown length
2337 2536: contig of 1320 bp in length
2637 2636: gap of unknown length
4058 4157: contig of 1421 bp in length
4158 5245: contig of 1088 bp in length
5246 5345: gap of unknown length
5346 6461: contig of 1116 bp in length
6462 6561: gap of unknown length
6562 7725: contig of 1164 bp in length
7726 7825: gap of unknown length
9144 9243: contig of 1318 bp in length
9244 10492: contig of 1249 bp in length
10493 10592: gap of unknown length
11667 11766: gap of 1074 bp in length
11767 12950: contig of 1184 bp in length
12951 13050: gap of unknown length
13051 14385: contig of 1235 bp in length
14386 15478: contig of 1093 bp in length
15479 15578: gap of unknown length
15579 16779: contig of 1201 bp in length
16780 16879: gap of unknown length
16880 18278: contig of 1399 bp in length
18279 18378: gap of unknown length
18379 19452: contig of 1074 bp in length
19453 19552: gap of unknown length
19553 21012: contig of 1460 bp in length
21013 21112: gap of unknown length
21113 22169: contig of 1057 bp in length
22170 22269: gap of unknown length
22270 23367: contig of 1098 bp in length
23368 23467: gap of unknown length
23468 24559: contig of 1092 bp in length
24560 24659: gap of unknown length
24660 25825: contig of 1166 bp in length
25826 25925: gap of unknown length
25926 27036: contig of 1111 bp in length
27037 27136: gap of unknown length
27137 28245: contig of 1109 bp in length
28246 28345: gap of unknown length
28346 29657: contig of 1311 bp in length
29657 30883: gap of unknown length
30884 30983: contig of 1127 bp in length
30984 32125: contig of 1142 bp in length
32126 32225: gap of unknown length
32226 33397: contig of 1172 bp in length
33398 33497: gap of unknown length
33498 34645: contig of 1148 bp in length
34646 34745: gap of unknown length
34746 35894: contig of 1149 bp in length
35895 35994: gap of unknown length
35995 37156: contig of 1162 bp in length
37157 37256: gap of unknown length
37257 38396: contig of 1140 bp in length
38397 38496: gap of unknown length
38497 39570: contig of 1074 bp in length
39571 39670: gap of unknown length
39671 40852: contig of 1182 bp in length
40853 40952: gap of unknown length
40953 42116: contig of 1164 bp in length
42117 42216: gap of unknown length
42217 43372: contig of 1156 bp in length
43373 43472: gap of unknown length
43473 44603: contig of 1131 bp in length
44604 44703: gap of unknown length
44704 45924: contig of 1221 bp in length
45925 46024: gap of unknown length
46025 47174: contig of 1150 bp in length
47175 47274: gap of unknown length
47275 48744: contig of 1470 bp in length
48745 48844: gap of unknown length
48845 49969: contig of 1125 bp in length
49970 50069: gap of unknown length
50070 51135: contig of 1066 bp in length
51136 51235: gap of unknown length
51236 52690: contig of 1455 bp in length
52691 52790: gap of unknown length
52791 53920: contig of 1130 bp in length
53921 54020: gap of unknown length
54021 55101: contig of 1081 bp in length
55102 55201: gap of unknown length
55202 56324: contig of 1123 bp in length
56325 56424: gap of unknown length
56425 57583: contig of 1159 bp in length
57584 57683: gap of unknown length
57684 58822: contig of 1139 bp in length
58823 58922: gap of unknown length
58923 60547: contig of 1625 bp in length
60548 60647: gap of unknown length
60648 61743: contig of 1096 bp in length
61744 61843: gap of unknown length

```


REFERENCE
AUTHORS
TITLE
JOURNAL

MO 63108, USA
5 (bases 1 to 196204)
Waterston,R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 196204)
Waterston,R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 29, 2001 this sequence version replaced gi:14589737.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics

Center project name: H_NH0449P15

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mail to: egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-1244M4; the clone sequenced to the right is RP13-580H13. Actual start of this clone is at base position 1 of RP11-449P15; actual end is at base position 196204 of RP11-449P15.

Data from AC091729 was used to finish this clone, AC073957.
Polymorphisms have been identified between AC073957 and AC091729.

Unresolved simple sequence repeats, base positions 26410 to 27787 and 63334 to 64953.

FEATURES
source

Location/Qualifiers
1..196204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"

/clone="RP11-449P15"
/clone_lib="RPCI-11"
1..278
/note="match to EST BG992151 (NID:g14396221)"
203..598
/note="match to EST BF814714 (NID:g12147717)"
1010..1156
/note="match to EST F06563 (NID:g672150)"
1122..1414
/rpt_family="Alu"
1344..1863
/note="similar to Homo sapiens EST BF928866 (NID:g12326994)"
1932..2224
/rpt_family="Alu"
2189..2393
/note="match to EST F06563 (NID:g672150)"
2270..2636
/rpt_family="L1"
2393..2422
/rpt_family=" (CACCC)n"
2745..3367
/rpt_family="L1"
4057..4262
/note="match to EST BI039739 (NID:g14446365)"
5018..5019
/note="match to EST AW877271 (NID:g8015626)"
5068..5455
/note="match to EST BF087410 (NID:g10893120)"
5193..5455
/note="match to EST AA004851 (NID:g14485598) zh92f02.rl"
5405..5761
/note="match to EST AW877271 (NID:g8015626)"
5829..5833
/note="match to EST BG992151 (NID:g14396221)"
5844..6140
/rpt_family="Alu"
6107..6377
/note="match to EST BE073441 (NID:g8420427)"
6136..6426
/note="match to EST BF874571 (NID:g12264701)"
6174..6844
/notes="match to EST AUI33804 (NID:g10994343)"
7097..7195
/note="similar to Mus musculus EST AA021831 (NID:g1485585) mh82C08.rl"
7108..7195
/note="match to EST AI200352 (NID:g3752958) qf92h12.xl"
7108..7195
/notes="match to EST BF516594 (NID:g11601773)"
7108..7195
/notes="match to EST BG709448 (NID:g13987794)"
7108..7195
/note="match to EST T16043 (NID:g518205)"
7109..7195
/note="match to EST AL527751 (NID:g12791244)"
7117..7199
/note="match to EST AA004851 (NID:g1448598) zh92f02.rl"
7117..7195
/note="match to EST BF087410 (NID:g10893120)"
7141..7195
/note="match to EST AI174672 (NID:g6361050)"
7150..7195
/note="match to EST BG034421 (NID:g12427714)"
7628..7711
/note="similar to Mus musculus EST W65144 (NID:g1372795) me02f03.rl"
7630..7711
/note="match to EST AI174672 (NID:g6361050)"
7630..7711
/note="match to EST AI200352 (NID:g3752958) qf92h12.xl"
7630..7711
/note="match to EST AL527751 (NID:g12791244)"

Summary statistics
Consensus quality: 71349 bases at least Q40

Db
26974 AAGGAGAAAGGGAGATAGGAGGAATAGGGAGAGAGAGATTACCGCTTTCCTTT

* NOTE: This is a 'working draft' sequence. It currently
the following cutoffs: length >= 200 bases.

```

* 153671 154319: contig of 649 bp in length
* 154320 154399: gap of unknown length
* 154743 154743: contig of 344 bp in length
* 154744 154823: gap of unknown length
* 154824 154922: contig of 599 bp in length
* 155423 155502: gap of unknown length
* 155503 155840: contig of 338 bp in length.
FEATURES
    Location/Qualifiers
        1..155840
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /strain="y; cn bw sp"
            /db_xref="taxon:7227"
            /chromosome="3"
            /map="90E1-90F10"
            /clone="BACR48113 (D522) RPCI-98 48.1.13"
            /clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6"
ORIGIN
Query Match      7.2%; Score 97.4; DB 2; Length 155840;
Best Local Similarity 45.5%; Pred. No. 1.1e-10;
Matches 351; Conservative 0; Mismatches 409; Indels 11; Gaps 2;
QY 279 GCAGGAGGAGCGACAAAGAAAGCAACAGAAACCGCCGACCGAGCGGACCAACGAGAG 338
Db 135448 GGAGGACAAAGGGGAAACAGAGAAAGNGAACAAGGGGGGAGGGGGGGGAGCGGNGAG 135389
QY 339 CAGAGACCCCGAGCGCCCG---GAGAGCGAGAGAGCGACACACACACAGAGGGCGCCCG 395
Db 135388 CAGCAGCAGGGCGCAACANAGAGGGAACAAGANAGNGGCGACACACGCGGAGCAGGNG 135329
QY 396 GCGAAGCAGGAAACAGAAACACCGAGGCGAGGACACACACACAGAGAGCGCCCGAGAG 455
Db 135328 CGGGCGGCGACAAAGCGGGGAGACACGAGAGAACGCAACAGAGGNGAGGAGGANAAGA 135269
QY 456 AGGAGAGCAGCGGGAGCAGAGAGGAGGAGCGAGGAGCGAGGAGGAGGAGGAGGAGGAG 515
Db 135268 CNGCGGNGCAGCAGAGGAGCAACACACACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135209
QY 516 GCGCCACACACACCGCGGAGAGAGAGAAACCAAGAGGCGCAGAGGAGCGACCAAGAGAG 575
Db 135208 GGAGGAAAGAGAGAGCAGGCGCANAGCAGCGCCAGAGCGGGAANGAGGGGCAAGAGCG 135149
QY 576 CGAGAAAGGGCGCGAGAGAGAAACCGCGCA-----ACAGCGGCGAGAGCAAAA 627
Db 135148 GGACAAAGGGCGCGAGGCGGCGGAGCCAGCGCGGAGAGAGACGAGGGGACAGAN 135089
QY 628 AGCAGAGCCAAAGAGCAGCAAAAGACCGAAGACCGCCAGAGCGCCCAAGAGCGCCACCA 687
Db 135088 GGCAAGGCAACAGGCAACCCANAGGAGAGCGGAGCGGAGAGAAAGAGCGCCAGCG 135029
QY 688 AGAAGAGCAGGCGAGAGAGAAACCAACACAGCGCGGAGAGCGGAGCGCGGCGCGA 747
Db 135028 GGGCAGCGCAGGCGANAAAGAGNAGCAAAAGACAGCGCCCAAAAGCGCCAGCG 134969
QY 748 GAGGAGAAACACCCACCGCGGAGCCACAAAGAGACCCCGCCCAACACAGAGAGAG 807
Db 134968 GAGGAGAAACCGAGAGCAGCAGCAGCCANACAGAGCAACACGAAAGAGAGAGGGGG 134909
QY 808 GAGAACACCGAGAGCGAAAGAGAGGAACAAGCGGCGAGAGCGGAGCGAGCGCGCGCG 867
Db 134908 CAAAGAACCAACAGGCGCAAAAGAGAGCGCGCGGCGCANAGAGANACAAAGAN 134849
QY 868 CCACAGAGCAGAGAAAGCAGAGAGACCAAGAAAGAGCGCGGAGAGAGAGCGCA 927
Db 134848 AACCGGAGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134789
QY 928 GCAGAAAGGAGGAGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
Db 134788 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134729

```

```

QY 988 GGGCAACACAGAACAAAGAACGGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1038
Db 134728 AAACCCGAGAGCGGAGCAGACACCAAGACAGAGGAGAGAGAGAGAGAGAGAGAG 134678

RESULT 15
AC147269/c
LOCUS AC147269 177277 bp DNA linear ROD 15-MAY-2004
DEFINITION Mus musculus BAC clone RP24-281E6 from chromosome 18, complete
sequence.
AC147269
AC147269.3 GI:45237303
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Berghoff,A., Haglund,K. and Bielicki,I.
TITLE The sequence of Mus musculus BAC clone RP24-281E6
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Mar 6, 2004 this sequence version replaced gi:39979572.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
-----
Center project name: M_BR0281E06

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be

Sat Oct 23 17:44:55 2004

obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is overlapped by AC139334.

Location/Qualifiers

1. .177277
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"
/clone="RP24-281B6"
/clone_lib="RPC1-24"
105840..105910
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=32.06 / Sec struct Sc=-10.06)"

tRNA

ORIGIN

Query Match 7.2%; Score 97.2; DB 10; Length 177277;
Best Local Similarity 46.8%; Pred. No. 1.2e-10;
Matches 378; Conservative 0; Mismatches 423; Indels 7; Gaps 2;
QY 398 GAAGCAGGAACCCAGAAACCGGGCGGACCAACCCGAAAGGAGGACCCAGGAGACAG 457
Db 97498 GAAGGAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 97439
QY 458 GAGAAAGCAGCGGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 517
Db 97438 GGAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 97381
QY 518 CCCACACACGACCCGCGAAGAAAGAAACCAAGAAAGGCGCACAGGGCAGCCAGGAAGCG 577
Db 97380 CAGACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97321
QY 578 AGAAAGGGCCGCGAGGAAGAAGAACCGCGCAACACAGGGCAGAGACAAAGACAGGCCA 637
Db 97320 AAAGAAAGACACAGCAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97261
QY 638 AAAGAAAGCAGCAAAAGACCGAGAGCCGCGAGAGCCACAAAGCCCAACCAAGAAAGACAC 697
Db 97260 GACAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97201
QY 698 GGGACGAGAGGAAACCCCAACACAGCCGAGACCGAGCCCGGCGCCGAGAGGAGAAC 757
Db 97200 AGAAAGAAAGAAAGACAGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97141
QY 758 CCCACACCCCGAAGCCACAAAGAACCCACCCCAACACAGAGGAGAGGAGAACCA 817
Db 97140 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 97086
QY 818 GGNAGCCAAAGNAGAGGACACAGCCGCGAGAGCGGAGCAGCCCGCGCCACAGAGA 877
Db 97085 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 97026
QY 878 CAGAGAAAGACAGCAGAGACCAAGAAAGGAGCCGCGAGAGAAAGCCAGCAGAAAGGA 937
Db 97025 AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGG 96966
QY 938 AAGGCACCGCGGAGGAAGCCAGACAGCCGCAAGAGAGGAGGAGGAGGAGGAGGAGG 997
Db 96965 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 96906
QY 998 GAACAAAGACGGCAGGAACCGAAGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1057
Db 96905 GAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 96846
QY 1058 GGCAGAAAGGGACCGGAAAGACCCAGCCACCCGAGAGGAGGAGGAGGAGGAGGAGG 1117
Db 96845 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGG 96786
QY 1118 CAACAAAGGGGAAGGGCCAAAGACACACCCCAACGAAAGACAGGGACAGAAAGGGCGC 1177

Db 96785 AGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAG 96726
QY 1178 CGAGAAAGAAAGAGGAGCGCAGAGAAAA 1205
Db 96725 AGAGAGAGAGAAATGGAAAGCCATGGACA 96698

Search completed: October 21, 2004, 05:53:03
Job time : 5820 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2004, 17:04:23 ; Search time 728 Seconds
(without alignments)
9691.239 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599

Perfect score: 1344
Sequence: 1 aagacacagcaggacacgc.....gagaagagagaagaccccg 1344

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1344	100.0	2565	4	Aaf28296
2	103	7.7	10732	3	AA10594
3	102	7.6	6755	2	AAV21511
C 4	97	7.2	42999	8	ABX11086
C 5	97	7.2	114955	2	AA53491
C 6	94.8	7.1	41907	9	ADA37416
C 7	94.8	7.1	42998	8	AD14747
C 8	94.8	7.1	42999	6	ABS65032
C 9	94.8	7.1	42999	10	AA61411
10	91.2	6.8	6741	3	AA10595
C 11	89	6.6	5059	2	AA84332
C 12	88.8	6.6	2318	5	AD163176
C 13	88.8	6.6	3662	4	ABL05274
C 14	88.6	6.6	1200	6	ABQ39210
C 15	88.6	6.6	1200	6	ABQ39211
C 16	87.6	6.5	1840	10	ADC87550
C 17	84.8	6.3	1762	6	ABQ20939
C 18	84.8	6.3	1762	6	ABQ20938
C 19	83.2	6.2	42533	8	AA856114
C 20	83.2	6.2	42533	9	ADA02476
C 21	83.2	6.2	42533	10	ADB72215

C 22	81.8	6.1	7033	3	AAA34722	Aaa34722 Human ade
C 23	81.8	6.1	7033	3	AAF20844	Aaf20844 Human low
C 24	81.8	6.1	7033	10	ABZ96538	Abz96538 Human nuc
C 25	81.8	6.1	7036	2	AA55274	Aax55274 Human fac
C 26	81.8	6.1	209273	3	AAF21437	Aaf21437 Human fac
C 27	81.8	6.1	209274	10	ABZ97131	Abz97131 Human enz
C 28	81.4	6.1	6668	6	ABL33697	Ab133697 Human imm
C 29	81	6.0	1839	4	AAH17145	Aah17145 Human cdn
C 30	80.6	6.0	901	6	ABQ13742	Abq13742 Oligonuc1
C 31	80.6	6.0	901	6	ABQ13743	Abq13743 Oligonuc1
C 32	80.6	6.0	1343	6	ABZ78056	Abz78056 Human bre
C 33	80.6	6.0	1343	6	ABZ78057	Abz78057 Human bre
C 34	80.4	6.0	1236	3	AAA02163	Aaa02163 Human col
C 35	79.4	5.9	887	8	ABT42734	Abt42734 Human neu
C 36	79.4	5.9	887	8	ABQ20942	Abq20942 Oligonuc1
C 37	78.8	5.9	1995	6	ABQ20943	Abq20943 Oligonuc1
C 38	78.8	5.9	1995	6	AAI94064	Aai94064 Human neu
C 39	78.4	5.8	863	4	AAI94286	Aai94286 Human neu
C 40	78.4	5.8	863	4	ADL42751	Adl42751 Human neu
C 41	77.2	5.7	1041	5	ADL40874	Adl40874 Human ova
C 42	77	5.7	876	6	ABQ75278	Abq75278 Human lun
C 43	77	5.7	1186	10	ADC87306	Adc87306 Human gpc
C 44	77	5.7	1225	10	ADC87492	Adc87492 Human gpc
C 45	76.2	5.7	5242	5	ABV24723	Abv24723 Human pro

ALIGNMENTS

RESULT 1
AAF28296
ID AAF28296 standard; RNA; 2565 BP.
XX
AC AAF28296;
XX
DT 05-APR-2001 (first entry)
XX
DE LPL mRNA.
XX
KW Lipoprotein lipase; LPL; S447X; cardiovascular; diabetes; obesity;
KW deficiency; ss.
XX
OS Homo sapiens.
XX
PN WO200100220-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-CA000762.
XX
PR 24-JUN-1999; 99EP-00202048.
XX
(UYBR-) UNIV BRITISH COLUMBIA.
(AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
(UYAM-) UNIV AMSTERDAM ACADEMIC HOSPITAL.

Hayden MR, Kastelein JJP;
WPI; 2001-112388/12.

Use of lipoprotein lipase S447X therapeutic for the preparation of a pharmaceutical composition for the treatment of coronary heart disease, atherosclerosis, angina pectoris, hypertension, diabetes, cachexia and obesity.
Claim 5; Fig 4; 5lpp; English.

The present invention relates to use of a lipoprotein lipase (LPL) S447X therapeutic for the preparation of a pharmaceutical for the treatment of an LPL-responsive condition. Diseases that may be treated are complete LPL deficiency, chylomicronemia, hyperlipidemia, partial LPL deficiency, pancreatitis, hypertriglyceridemia, hypoalbuminemia (low high density lipoprotein-cholesterol), cardiovascular disease, coronary heart

Best Local Similarity 15.3%; Pred. No. 2.7e-12;
Matches 207; Conservative 522; Mismatches 611; Indels 12; Gaps 4;

QY	4	ACACAGGAGGACACGCCACCCACCGGAGAGAGAGCGCGGACCGCACAACACAGCAG 63		Db	10135	GNHASTRASGTHRAGSRASGYCGASNGTHRHVSAGYSGARGHSRARGGASNSRRGNV 10194
Db	9115	ASRYSGHGYSGAGMTVAGGNSRGTHRSRYGARGAAYSAAGVAASASNSRHSNYSR 9174		QY	1073	AAAAACCCAGCCACCGGAGAGACACCCAGCACAACAAAAGGCGCAACAAAGGGAAGG 1132
QY	64	CAAAACCCAGGACACGAGCGGACGGAACAGGAAGAGAGAGGCGGCCAAAACCGCGGCC 123		Db	10195	ATHRTHRSRCYSGYHSASNGSNVSGTGYTHYRASNGYSGHTRASYSASYSRAAGSRGN 10254
Db	9175	HAAARGGASNGVARGAGGNGTHRGNAAGGASHGSHYAAMTSRARGGSRMTGYASANT 9234		QY	1133	GCCAAAAGCACACCCCAACGAAAGACAGGACAGAGAAAGCGGCGGAGAAAAGAAAGG 1192
QY	124	GACAGAGAACACAGACCAAGCAGGGGACGGCGACCGGCCAGGAGCAACCCAGGCC 183		Db	10255	YSGVAGARGYSGNSRAATHRASARGGSRMT-ASNGNGNGSRGYSTRASSRARGGNSRH 10313
Db	9235	HRSRTHRTYRGVAARGVAGYSGARGASNGGARGGHSRARGARGTHRAAGVARSYSYGN 9294		QY	1193	AGCGCAGAGAAAAGGCGGCCAGGCGGAGGCGCAAGCCAAAGAGAAACCCCAAGGAAAG 1252
QY	184	CGCGCACACCAACCGGGGACAGAGGCGCGGACAGGAGGAGGAGGAGGAGGAGGAGG 243		Db	10314	YSGSRASNGYNGASNGYSASNGNYSASNSRASNGHSTRGAAGYSAAAHASNSY 10373
Db	9295	NGSRTHRYSYNAYSGCYSGNARGVAAARGYSARGRGYRAAAAAAYSMYASNGV 9354		QY	1253	GACCGGACGAGCGCCCAACACCAAGCCAGGACCAAGAAAGCCGCGGCGGACGACAGCA 1312
QY	244	GGACAAAGCCCGGGAACAGCGGAGCCCGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 9354		Db	10374	SVASRGGGASNYASNSRASNCYVGGAAYSCYSGNGNGSRMTSYSGCYSSRYSHA 10433
Db	9355	AGMTGYARGASRHGARGARGSRYSRTHRASNSRVAGSRVAASTHRSGTHRRTHR 9414		QY	1313	CCACACAGCAG 1344
QY	304	CAACAGAACCGCCGACGAGGACCAACAGGAGCAGAGAGAGAGAGAGAGAGAGAGAG 9474		Db	10434	SASYSASYSRARGASNASTRGTHRAAAASRGY 10465
Db	9415	YSASNTHRTHRGNCYSAAVAGGASNYSAVSGSRASNGYASNGNSRARGVAMT 9474				
QY	364	AGAGAGCGCACACACACAGAGGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 421		RESULT 3		
Db	9475	THRARGTHRASRYSHSGSRNGASHSYSGVHSSRVAGHSHSRASNSRASNSRHSRH 9534		AAV21511		standard; DNA; 6755 BP.
QY	422	--CAGGACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479		XX	AAV21511;	
Db	9535	AASRSRAAGYASASYSYTHRSYASAGSRTAASRAASRASHSHARGSGYSY 9594		DT	17-OCT-2003 (revised)	
QY	480	GAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539		DT	17-AUG-1998 (first entry)	
Db	9595	SGRSYCSYSSRVAGYAAASASASNTASASHVAGMTGYSAAVASRVAGYSAATHRG 9654		XX	Staphylococcal bacteriocin BacRI operon.	
QY	540	AGAAACCAAGGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599		XX	BacRI; bacteriocin; antimicrobial; antibacterial; antibiotic;	
Db	9655	SRSRASRYSAVATHRGASNSRSGTHRYTHRYSGSRTHRGGYSYSGVARSASHS 9714		XX	Moraxella bovis; infectious bovine keratoconjunctivitis; cancer; therapy;	
QY	600	AACCGCGCAACACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 659		XX	ds.	
Db	9715	RGNTHRTHRTHRRGVAASMTGRTMSRASHSBSASRYSSRYASNTHRCTYSRA 9774		OS	Staphylococcus aureus; strain UT0007 (ATCC 55800).	
QY	660	GACCGCAGAGCCCAACAGCCCAACAGAGACACAGGAGGAGGAGGAGGAGGAGGAGGAG 719		XX	WO9812319-A1.	
Db	9775	SHYSGYTHRVARGYTRGNASVAVAYSTRGNGASNSHVSATHRSYSGYRASASHSASYSV 9834		PD	26-MAR-1998.	
QY	720	ACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779		PF	18-SEP-1997; 97WO-US016758.	
Db	9835	AAAARGHVAASNAASHASHASSRSRARGGYRVASHASTHRGNASNRGNCYSHSCY 9894		PR	19-SEP-1996; 96US-00710561.	
QY	780	AGACACCCCAACACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 839		PR	17-SEP-1997; 97US-00931999.	
Db	9895	SSRCYSASRSRVAASNSRRAAGYASNAASRYASRGSRGNGASRYSSRY 9954		XX	(UNIV) UNIV KANSAS STATE RES FOUND.	
QY	840	CAGC---GGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 895		XX	Iandolo JJ, Crupper SS;	
Db	9955	YSGVAGARGAAVAASTYRASSRSASRNGYSGYVARGASHSTRYVSASNGYMTTRTHR 10014		XX	WPI; 1998-230316/20.	
QY	896	ACCAGAAAAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 952		XX	Therapeutic proteinaceous substances from Staphylococcus aureus - useful	
Db	10015	GTYRMTVAARGVAGHNTYSTRSGRSYSTRSGRSYSTRSGRSYSTRSGRSYSTRSGRSY 10074		PT	to inhibit growth of wide range of prokaryotic or eukaryotic cells, e.g.	
QY	953	AAAGCAGACAAAGCCGAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012		PT	Moraxella bovis causing infectious bovine keratoconjunctivitis.	
Db	10075	YRGVASNHAAGVTHRTHRAAASSTRMTASHNSHCYSRSGNSVASHSRMTARGASAYSY 10134		PS	Claim 2; Page 19-23; 38pp; English.	
QY	1013	GGAACGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072		CC	This polynucleotide comprises the bacteriocin BacRI operon of	
				CC	Staphylococcus aureus UT0007. The sequence of the BacRI operon was	
				CC	determined by N-terminal sequencing of purified BacRI peptide (see	
				CC	AAW54171), with back-translation and plasmid analysis. The BacRI operon	
				CC	includes the BacRI gene (see AAV21510), a homologue of the cym gene of	
				CC	the cytotoxin operon of Enterococcus faecalis whose function is involved	
				CC	in the maturation of pre-cytotoxin, an ATP-transporter gene, bio1 and	
				CC	bio2 genes related to lactococin biosynthesis and modification, and a	
				CC	gene involved in immunity function. BacRI peptides can be produced by	
				CC	construction of an expression vector containing an oligonucleotide or	
				CC	operon coding for BacRI, and use of the vector to transform host cells	
				CC	for BacRI expression. The entire BacRI operon has been cloned into	

CC plasmid pUB110, and Bacillus subtilis transformants secreted the
CC recombinant BacRI peptide into the medium. Bacteriocin BacRI is active
CC against many Gram-positive and Gram-negative organisms such as Bordetella
CC bronchoseptica, Pasteurella multocida and Staphylococcus aureus;
CC Moraxella bovis, causing infectious bovine keratoconjunctivitis, is
CC especially sensitive. BacRI can also be used as an anti-cancer agent.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 6755 BP; 3903 A; 1252 C; 1600 G; 0 T; 0 U; 0 Other;

Query Match 7.6%; Score 102; DB 2; Length 6755;
Best Local Similarity 45.3%; Pred. No. 4e-12;
Matches 568; Conservative 0; Mismatches 675; Indels 11; Gaps 5;

QY 52 ACACACAGCAGCAGCAAAACCCAGGACCCAGGCGGACGAGCAAGCAAGGAGAGAGAGGGGCGCAA 111
DB 5413 ACAAGGGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5472
QY 112 AACGGGCGCGCGACAGAGAGAACACACCCAGAGCGGGGAGCGCGCAGCGGCGAGGA 171
DB 5473 AAAGGGCAAGCGAAGAGAGGAGAAACAGAAACCAAGCGACAGAGAGAAAGAAAGAACGGGAAC 5532
QY 172 GCAACCCAGCGCGCGGACACCAACCGGGGACAGGAGGCGCGGACACCGAGGAGGA 231
DB 5533 CCACAAGCAGGGAGGGGAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5592
QY 232 GGAGAACACCCCGGACAGGACCCAGCGGAAACAGCCGCGGAGCCCGCGCGCAGCAGGAAGCG 291
DB 5593 GGGGCAACACAGACAGCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5652
QY 292 ACCAAAGAAAGCAACACGCGGACCGGAGGAGGACCGAGCGAGAGAACAGGAGAGAGCGCGG 351
DB 5653 ACACAAAGAAAGCAACGCAAA--AACCCAGAACACGACAAAGAGGAGAAAGAAAGAAAG 5710
QY 352 CGCCCCGAGACAGAGACGACACACACACACGAGGCGGCGCGCGCGCGCGGAGCAACCG 411
DB 5711 ACAACCAAGACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5770
QY 412 AAACCGAGGCGGAGCAACCGGAGGAGGACCGAGCGAGAGAACAGGAGAGAGACCGCGG 471
DB 5771 CAGGCGCGCGGACAGAAAG 5830
QY 472 AGCAG 531
DB 5831 ACACCCAGAAACACCAAC 5890
QY 532 CCGGAGAGAGAAACCAAGAGCGGACACAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 591
DB 5891 AGGGAAGAGGAAACCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5950
QY 592 GGAGAAAGAACCGCGCAACACCGGCGGAGAGACAAAGAACAGAGCCCAAGAAAGAGAGCGCAA 651
DB 5951 AGAAGGGGAAAGAAACACGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6010
QY 652 AGACCGAAGACCGCGAGCGGACCAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
DB 6011 ACGAGAAAGAAAG 6070
QY 711 --ACCCACCAACAGCGCGGAGCGGAGGACCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
DB 6071 GGAG 6130
QY 769 GAAGCCACAAAGAGACCCCAAGACACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
DB 6131 GAAGGAG 6190
QY 829 GAAGAGAGACACAGCGGAG 888
DB 6191 ACCAAAGAAAG 6249
QY 889 AGGAG 946
DB 6250 AAAAAACGAGGAG 6309

QY 947 CGGAGGAGAGCCAGACAGCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
DB 6310 CACAAAG 6369
QY 1007 ACGGCGAGGAG 1066
DB 6370 AACAAAG 6429
QY 1067 GACCGAAAGAACCCAGCCCGAG 1126
DB 6430 AAAAAACCAACGAGCGGAG 6489
QY 1127 GGAAGGCGGCAAGAGCAGCCCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186
DB 6490 ACAGCAACAAAG 6546
QY 1187 AAAAGGAGCGCGAG 1246
DB 6547 GAGAACAG 6606
QY 1247 GBAACGAGCGGAG 1300
DB 6607 AACACGAG 6660

RESULT 4
ABX11086/c
ID ABX11086 standard; DNA; 42999 BP.
XX
AC ABX11086;
XX
DT 24-APR-2003 (first entry)
XX
DE Human ribosomal RNA (rRNA) gene.
XX
KW Nucleic acid amplification; artificial chromosome isolation; MAC;
KW DNA delivery; mammalian artificial chromosome; gene therapy; organ;
KW humanised genetically transformed animal; chromosomal element;
KW gene product production system; transgenic; centromere function;
KW information storage vehicle; artificial chromosome vector; human;
KW species-specific artificial chromosome; ribosomal RNA; rRNA; gene; ds.
XX
OS Homo sapiens.
XX
PN US2002160410-A1.
XX
PD 31-OCT-2002.
XX
PP 17-APR-2002; 2002US-00125767.
XX
PR 10-APR-1996; 96US-00629822.
PR 15-JUL-1996; 96US-00682080.
PR 07-AUG-1996; 96US-00695191.
PR 10-APR-1997; 97US-00835682.
PR 28-NOV-2000; 2000US-00724693.
XX
PA (HADI/) HADIACZY G.
PA (SZAL/) SZALAY A. A.
XX
PI Hadlaczky G, Szalay AA;
XX WPI; 2003-265757/26.
XX
DR Amplifying nucleic acid for constructing artificial chromosomes,
XX PT comprises introducing a DNA fragment into a cell to incorporate the DNA
XX PT into a chromosome, and identifying cells with chromosomes that has
XX PT undergone amplification.
XX
PS Disclosure; Page 76-94; 107pp; English.
XX
CC The present invention relates to methods for amplifying nucleic acids in
CC cells, for isolating artificial chromosomes, and preparing cell lines

25442	CAGAGAAACACAGAGAGAAAGAGAAAGACACAGAGAGCGGAGAGAGGGAGAGACAGAGAA	25387
687	AAGAAGACACGGGACAGAGAGGAAACCCAAACCAACAGCCGCGACACCGAGGCACCGGGCCG	746
25382	CAGAAAGCGCGGAGAGAGACACAGAGAGACACAGACAGATAGACAGGCAGAGAAAGAGAGT	25323
747	AGAGGAGAAACCCACACCCGCCGAAGCCACAAAGAACCACCCCAACACACAGAGAGAA	806
25322	AAGCAGAGATAGGCACAGAGAGAGAGACAGAGAGACACAGAGAAAGAGAGAGAGGC	25263
807	GGAGAACACACAGGAAGCCAAAGGAAGGACACAGCGGCAGACGGGGGACAGCCCGGCC	866
25262	AGACAGACAGAAAGGGACACAGACAGAAAGACAGAGACAGAGAGAGAGAGAGAGAGAG	25213
867	GCCACAGAGACAGAGAAAGACAGAGAGACCGAGAAAGGGACCGCAGGAGAGAAAGGCC	926
25212	-----AGAGAGAGAGAGAGAGAGAAACAGACAGAAAGAGAGAGAGAGAGAGAGAGAG	25159
927	AGCAGAAAGGAAAGSCACCGCGAGGAAGCCAGACAAAGCCGCGGAAAGAGAGAGAGAG	986
25158	AGACAGACAGACAGACGACAGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	25099
987	GGCGGAAACACAGAACAAAGACAGGACGGAACCGGAAGAGAGAGAGAGAGAGAGAGAG	1046
25098	CGGACAGACACAGAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG	25039
1047	AACCCAGGGGGGCAAAAGGACCGGAAACCCAGCCACCCGAGAGAGGAGACAGCCAAAG	1106
25038	GAGAGAGACATGAGAGAAACACAGACAGAAAGAGGGAGACAGAGAGAGAGAGAGAG	24979
1107	CACAAAAAGGGCAACAAAGGGGAAAGGGCCAAAGCACACCCCAACGAAGACAGAGGACA	1166
24978	GNNNACAGAG-----AGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	24926
1167	GAAAGGGCGCCGAGAGAAAGAGAGCGGACAGAGAAAGAGGCC-----CAGGGCGAGG	1223
24925	GACAGACAGACAGGACAGAGAAAGAGAGTAAGACAGAGATAGGACAGAGAGAGAGAG	24866
1224	CCAGCCAAAGGAGAACCCCAAGGAACGACCGGACGAGGCC-----CAAAACCAAGCC	1280
24865	ACAGACAGAGAGACACAGAAAGAAAGAGAGGGGCGACAGACAGAGAGAGAGAGAGAT	24806
1281	CAAGAACAGAAAGCCCGCGCACGAAACAGACCACACACAGACAGAGAGAGAGAGAG	1335
24805	AGAGAAAGAGAGAGGCGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	24751
RESULT 8		
ABS65032/c		
ID	ABS65032 standard; DNA; 42999 BP.	
XX	AC	
XX	ABS65032;	
XX	DT	
XX	15-NOV-2002 (first entry)	
XX	Invertebrate foraging behaviour associated human DNA sequence #12.	
XX	Attention deficit hyperactivity disorder; ADHD; hypertension;	
XX	invertebrate foraging behaviour; nitric oxide; hypotensive;	
XX	cGMP-dependent protein kinase; human; neuroleptic; ds; gene therapy.	
OS	Homo sapiens.	
XX	WO200259370-A2.	
XX	PD	
XX	01-AUG-2002.	
XX	13-DEC-2001; 2001WO-US048087.	
XX	15-DEC-2000; 2000US-00738630.	
XX	(NEUR-) NEUROSCIENCES RES FOUND INC.	

PI Greenspan RJ, Shaw PJ;
 XX WPI; 2002-636544/68.
 XX
 PT Identifying a compound that modulates an attention deficit hyperactivity
 PT disorder (ADHD) for treating e.g. ADHD or hypertension, comprises
 PT measuring a foraging behavior in an invertebrate.
 PS
 PS Claim 89; Page 182-195; 246pp; English.
 XX
 CC The present invention relates to a method of identifying a compound that
 CC modulates attention deficit hyperactivity disorder (ADHD) in a mammal.
 CC The method comprises administering a test compound to an invertebrate,
 CC and measuring a foraging behaviour of the invertebrate, where the
 CC compound that modulates the foraging behaviour of the invertebrate is
 CC characterised as a compound that modulates behaviour of the invertebrate is
 CC identified by the new method of the invention is useful for diagnosing or
 CC treating ADHD, hypertension or other diseases associated with a nitric
 CC oxide/cGMP-dependent protein kinase network in a mammal, particularly
 CC humans. The method distinguishes a compound that has a specific effect on
 CC ADHD, hypertension or other diseases associated with a nitric oxide/cGMP-
 CC dependent kinase protein network in a mammal from a compound that has a
 CC non-specific effect. The present sequence represents a DNA sequence
 CC associated with invertebrate foraging behaviour
 XX
 SQ Sequence 42999 BP; 6411 A; 13605 C; 11491 G; 11479 T; 0 U; 13 Other;
 Query Match 7.1%; Score 94.8; DB 6; Length 42999;
 Best Local Similarity 45.3%; Pred. No. 2e-10;
 Matches 568; Conservative 0; Mismatches 655; Indels 32; Gaps 5;
 *

QY 90 ACAGAAAGAGAGAGGGGGCCAAAACGGGGCCCGCCGACAAAGAGAGAACCCAGCCCAACG 149
 DB 25982 ATAGAAACAG 25923
 QY 150 GGGGACCGGACCGGGCCAGAGAGCAACCCAGCCCGGGCCACCAACCGGGGACACGGA 209
 DB 25922 ACAGACAGAGAAACAGACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25863
 QY 210 GGGCCCGGACACCGAG 269
 DB 25862 AGAAACAGACAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25803
 QY 270 CCAG---CGCGGACGAG 326
 DB 25802 GACAGACACAGTACAG 25743
 QY 327 GACCAAG 386
 DB 25742 GACAGAGAGAAAG 25683
 QY 387 AGGGCCCGGCGAAGCAG 446
 DB 25682 AGAGAGAGAAACAGGAG 25623
 QY 447 CAGGAGAAACAGGAG 506
 DB 25622 GAG 25563
 QY 507 CCCCAGAGGGCCACACCCACGACCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
 DB 25562 ACACAGTACAG 25503
 QY 567 CAAGGAAGCCGAG 626
 DB 25502 ACACAGAGAAAG 25443
 QY 627 AAGCAGAGCCAAAG 686
 DB 25442 CAG 25383
 QY 687 AAGAAAGACCGGAG 746

DB 25382 CAGAAAGCGCAGGAGT 25323
 QY 747 AGAGAGAGAAACACCCACACCCCGCGAAGCCACACAAAGAGACCACCCCAAAACACAGAGAGAGAA 806
 DB 25322 AAGACAGAGAGATAGGCACACAG 25263
 QY 807 GGAGAAACACCGAG 866
 DB 25262 AGACAG 25213
 QY 867 GCCACAG 926
 DB 25212 -----AG 25159
 QY 927 AGCAAGAAAGGAG 986
 DB 25158 AGACAG 25099
 QY 987 GGGCGAACACAG 1046
 DB 25098 GGGAG 25039
 QY 1047 AACCCAGGGGGGCAAAAG 1106
 DB 25038 GAG 24979
 QY 1107 CACAAAAGAGGAG 1166
 DB 24978 GNNAGAGAG-----AG 24926
 QY 1167 GAAAGCGCGCGAG 1223
 DB 24925 GACAGACAG 24866
 QY 1224 CCAAGCCCAAG 1280
 DB 24865 ACAGACAG 24806
 QY 1281 CAAGAAACAG 1335
 DB 24805 AG 24751

RESULT 9
 AAD61411/c
 ID AAD61411 standard; DNA; 42999 BP.
 XX
 AC AAD61411;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human ribosomal DNA (rDNA) repeat region.
 XX
 KW Satellite artificial chromosome; SATAC; transgenic; transgenic plant;
 KW centromere; human; ds.
 OS Homo sapiens.
 XX
 PN US2003101480-A1.
 XX
 PD 29-MAY-2003.
 XX
 PF 01-NOV-2002; 2002US-00287313.
 XX
 PR 10-APR-1996; 96US-00629822.
 PR 15-JUL-1996; 96US-00682080.
 PR 07-AUG-1996; 96US-00695191.
 PR 10-APR-1997; 97US-00835682.
 PR 28-NOV-2000; 2000US-00724726.
 XX
 PA (HADI/) HADLACZY G.
 PA (SZAL/) SZALAY A A.

PI Hadiaczky G, Szalay AA;
XX WPI; 2003-777325/73.
DR
XX Isolated substantially pure plant satellite artificial chromosome useful
PT for producing transgenic plant, producing gene product, cloning
PT centromere from plant.
XX
XX Example 6; Page 76-94; Opp; English.
PS
XX The present invention relates to novel pure plant satellite artificial
CC chromosome (SATAC). Sequences of the invention are useful for producing
CC transgenic plants and gene products. They are also useful for cloning a
CC centromere from a plant. The present sequence is human ribosomal DNA
CC (rDNA) repeat region used in the exemplification of the invention
XX
SQ Sequence 42999 BP; 6411 A; 13605 C; 11491 G; 11479 T; 0 U; 13 Other;
Query Match 7.1%; Score 94.8; DB 10; Length 42999;
Best Local Similarity 45.3%; Pred. No. 2e-10;
Matches 568; Conservative 0; Mismatches 655; Indels 32; Gaps 5;
QY 90 ACAGGAGAGAGGGGCGCCAAACCGGCGCGCCGCAACAGAGAGAACCCAGCCACAGCAG 149
DB ATAGGAACAG 25923
QY 150 GGGGACCGCGCCAGGAGCAACCCAGGCGCGCGGCAACCAACCGGCGGACAGGA 209
DB ACAG 25922
QY 210 GGGCGCGGACACCGGAG 269
DB AGAAACAGACAGTAAAG 25903
QY 270 CCCAG---CGCGGAG 326
DB GACAGACACAGTGTAG 25802
QY 327 GACCAACGAG 386
DB GACAG 25742
QY 387 AGGGCGCGGAG 446
DB AAG 25682
QY 447 CAGGAG 506
DB GAG 25622
QY 507 CCCCACGAGCGCCACACCCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
DB ACACAGTGTAG 25562
QY 567 CAAGGAGCGGAG 626
DB ACGACAG 25502
QY 627 AAGCAGAGCGAG 686
DB CAG 25442
QY 687 AAG 746
DB CAG 25382
QY 747 AAG 806
DB AAG 25322
QY 807 GAG 866
DB AAG 25262

QY 867 GCCCAG 926
DB 25212 -----AG 25159
QY 927 AGCAG 986
DB 25158 AGCAG 25099
QY 987 GGGCGAG 1046
DB 25098 GGGAG 25039
QY 1047 AACCCAGGGGGGCGCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
DB 25038 GAG 24979
QY 1107 CACAAAAGGGGCAACAAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
DB 24978 GNNAGAGAG-----AG 24926
QY 1167 GAAAGCGCGCGGAG 1223
DB 24925 GACAG 24866
QY 1224 CCAAGCCAAAGGAG 1280
DB 24865 ACAG 24806
QY 1281 CAAG 1335
DB 24805 AG 24751
RESULT 10
AAAL0595
ID AAAL0595 standard; DNA; 6741 BP.
XX
AC AAAL0595;
XX
DT 29-JUN-2000 (first entry)
XX
DE Gene encoding a subunit of cellulose synthase.
XX
KW Cellulose synthase; cellulose production; increase yield; ds.
XX
OS Vigna angularis.
XX
PN JP2000060568-A.
XX
PD 29-FEB-2000.
XX
PF 26-AUG-1998; 98JP-00239998.
XX
PR 26-AUG-1998; 98JP-00239998.
XX
PA (MIZU/) MIZUNO K.
XX
PA (OJIP) OJI PAPER CO.
XX
DR WPI; 2000-342371/30.
DR P-PSDB; AAY85180.
XX
PT A gene encoding a cellulose synthetic equipment - for the improvement in
PS the amount of cellulose synthesized in a plant body.
XX
PS Claim 2; Page 22-31; 32pp; Japanese.
XX
CC This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesized by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant

[illegible]

Db	6527	RMTRGVAAAGYGYMTHSRVARTYRASVAGYGYMTRVAARGAAGAAASRSRGNGNRGYAA
Qy	1309	CAGACACACACAGCAGAGAG 1330
Db	6587	SRAAAASNAASRRGNGNARG 6608
RESULT 11		
AXX84332/c		
ID	AXX84332	standard; DNA; 5059 BP.
XX	AC	AXX84332;
XX	DT	08-SEP-1999 (first entry)
XX	DE	Stealth virus nucleic acid clone, SEQ ID NO: 24.
XX	KW	Stealth virus; detection; diagnosis; infection; ss.
XX	OS	Stealth virus.
XX	Key	Location/Qualifiers
FH	misc_difference	3605
FT		/tag= a
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3610
FT		/tag= b
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3615
FT		/tag= c
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3630
FT		/tag= d
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3631
FT		/tag= e
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3638
FT		/tag= f
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3641
FT		/tag= g
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3642
FT		/tag= h
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3657
FT		/tag= i
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3659
FT		/tag= j
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"
FT	misc_difference	3681
FT		/tag= k
FT		/note= "this nucleotide is represented as a * in the
FT		specification, and is included to maintain the base
FT		numbering given in the specification"

XX	XX	Oligonucleotide for detecting cytosine methylation SEQ ID NO 25802.
DE	DE	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX	XX	drug; side effect; cancer; central nervous system; cardiovascular;
KW	KW	gastrointestinal; respiratory system; single nucleotide polymorphism;
KW	KW	SNP; cell differentiation; ds.
XX	XX	
OS	OS	Homo sapiens.
XX	XX	WO200218632-A2.
PN	PN	07-MAR-2002.
PD	PD	
XX	XX	
PF	PF	01-SEP-2001; 2001WO-EP010074.
XX	XX	
PR	PR	01-SEP-2000; 2000DE-01043826.
XX	XX	05-SEP-2000; 2000DE-01044543.
XX	XX	(EPIG-) EPIGENOMICS AG.
PA	PA	
PI	PI	Olek A, Pipenbrock C, Berlin K, Guetig D;
XX	XX	WPI; 2002-371829/40.
DR	DR	
XX	XX	Determining the degree of cytosine methylation in genomic DNA, useful for
PT	PT	diagnosis and prognosis, comprises selective hybridization of amplicons
PT	PT	from chemically treated DNA.
PT	PT	
XX	XX	Claim 12; 56pp + Sequence Listing; 56pp; German.
PS	PS	
XX	XX	This invention describes a novel method for determining the degree of
CC	CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC	CC	genomic sample of DNA. The sample is treated chemically to convert
CC	CC	cytosine (C) but not methylated C, to uracil, then part of the genomic
CC	CC	DNA that contains the target C is amplified to form a labeled amplicon.
CC	CC	The amplicon is hybridised to two classes, each with at least one member,
CC	CC	of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC	CC	degree of hybridisation to both classes is determined from the label on
CC	CC	the amplicon. From the ratio of labels hybridised to the two classes of
CC	CC	oligomers, the degree of methylation is calculated. The method is used:
CC	CC	(i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC	CC	and of a wide range of diseases, e.g. cancer, disorders of the central
CC	CC	nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC	CC	particularly by detecting mutations or single nucleotide polymorphisms
CC	CC	(SNP's); and (ii) for differentiation of cell or tissue types and for
CC	CC	investigating cell differentiation. The method allows the methylation
CC	CC	status of many C residues to be determined simultaneously. ABQ13410-
CC	CC	ABQ54121 represent genomic DNA sequences used to illustrate the method
CC	CC	for determining the degree of cytosine methylation described in the
CC	CC	disclosure of the invention
XX	XX	
SQ	SQ	Sequence 1200 BP; 919 A; 153 C; 112 G; 0 T; 0 U; 16 Other;
		Query March 6.6%; Score 88.6; DB 6; Length 1200;
		Best Local Similarity 43.8%; Pred. No. 2e-09; Indels 0; Gaps 0;
		Matches 349; Conservative 0; Mismatches 447;
Qy	536	AAGAAGAAAAACAAGAGGCCACAGCGGACGACCAAGGAGCCGAGAAAGGGCCGCGAGGAG 595
Db	192	AAGCGCGGCAAAAC 251
Qy	596	AAAGAACCGGCAACACCGGCGAGAGACAAAGCGAGCCCAAGAGACGACGCAAGAGAC 655
Db	252	GAAGAAAAACGCAAAAAAANPACAAANNAAAAAAAAAAAAAACGAAACGCGNCGAANA 311
Qy	656	CGAAGACCGCCAGAGGCCCAAGAGGCCAACCAAGAAAGACACGGGACGAGAGGAAACCCA 715
Db	312	CGAAACGAAACGAAAAAAAAAAGCNAACAAAAAACAACGACGCAACACGAAAAANA 371
Qy	716	ACCAACAGGCCGAGACCGAGGACCGGGCCGAGAGGAGAGACCCACACGCGCGAGGCCA 775
Db	372	AAAAAAAAAAAAAACAACGCGCAAAAAAACAACGAAAAAACAACAAAAAANAANA 431

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2004, 03:30:20 ; Search time 137 Seconds
(without alignments)
6972.995 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599
Perfect score: 1344
Sequence: 1 aagacacagcgagacacgc.....gagaagagagaagacccgc 1344

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	129.6	9.6	7218	1	US-08-232-463-14
C 2	102	7.6	6755	3	US-08-931-999-4
C 3	82.8	6.2	851	4	US-09-270-767-8286
C 4	82.8	6.2	851	4	US-09-270-767-8286
C 5	82.6	6.1	1827	4	US-09-270-767-23568
C 6	82.6	6.1	1827	4	US-09-270-767-1308
C 7	70.8	5.3	1298	3	US-08-948-705-3
C 8	70.8	5.3	1298	4	US-09-510-543-3
C 9	67.2	5.0	90541	4	US-09-759-359A-3
C 10	67.2	5.0	90541	4	US-10-207-973-3
C 11	63.8	4.7	53526	3	US-08-658-136-2
C 12	63.8	4.7	53577	3	US-08-658-136-1
C 13	63.6	4.7	705	4	US-09-270-767-5061
C 14	63.6	4.7	705	4	US-09-270-767-20343
C 15	63.6	4.7	1835	4	US-09-216-393B-80
C 16	63.4	4.7	1923	4	US-09-134-000C-466
C 17	62.4	4.6	1374	4	US-09-270-767-2575
C 18	62.4	4.6	1374	4	US-09-270-767-17857
C 19	61.2	4.6	176373	3	US-09-128-155-17
C 20	60	4.5	152331	3	US-09-128-155-16
C 21	59.6	4.4	1138	4	US-09-509-712B-107
C 22	59.4	4.4	289	3	US-09-007-005-17
C 23	59.4	4.4	289	3	US-09-244-796-17
C 24	56.6	4.2	152331	3	US-09-128-155-16
C 25	56	4.2	72604	3	US-09-268-992-7
C 26	56	4.2	72604	3	US-09-657-474-7
C 27	55.8	4.2	319608	4	US-09-539-333D-1

28	55.8	4.2	319608	4	US-09-679-409-1	Sequence 1, Appli
C 29	55.6	4.1	936	4	US-09-270-767-4464	Sequence 4464, Ap
C 30	55.6	4.1	936	4	US-09-270-767-19746	Sequence 19746, A
C 31	55.6	4.1	1559	3	US-09-019-095A-7	Sequence 7, Appli
C 32	54.2	4.0	510	4	US-09-248-796A-13735	Sequence 13735, A
C 33	54.2	4.0	168174	4	US-10-071-411A-63	Sequence 63, Appli
C 34	54.2	4.0	168174	4	US-10-071-411A-2	Sequence 2, Appli
C 35	54	4.0	435	4	US-09-328-352-583	Sequence 583, App
C 36	53.8	4.0	44453	3	US-09-146-053-5	Sequence 5, Appli
C 37	53.6	4.0	774	4	US-09-252-991A-3254	Sequence 3254, Ap
C 38	53.4	4.0	202001	4	US-09-734-674-3	Sequence 3, Appli
C 39	53	3.9	432	4	US-09-328-352-585	Sequence 585, App
C 40	53	3.9	447	4	US-09-328-352-584	Sequence 584, App
C 41	52.6	3.9	668	1	US-08-516-545-1	Sequence 1, Appli
C 42	52.6	3.9	668	4	US-09-050-344-1	Sequence 1, Appli
C 43	51.8	3.9	6158	4	US-09-799-451-897	Sequence 897, App
C 44	51.6	3.8	336	4	US-09-248-796A-13866	Sequence 13866, A
C 45	51.4	3.8	893	4	US-09-509-712B-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZpmt-Fls
US-08-232-463-14

Query Match 9.6%; Score 129.6; DB 1; Length 7218;


```
Best Local Similarity 1.6%; Pred. No. 6.1e-20;
Matches 6; Conservative 289; Mismatches 83; Indels 0; Gaps 0;
QY 870 ACAGACAGACAGAAAGCAGGAGAGACAGACAGAAAGGACCGCAGGAGAGAGAGCCAGC 929
Db 1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1377
QY 930 AGAAGGAAAGGACCGCGGAGGAGAAAGCCAGACAGAAAGGAGGAGGAGGAGGAGG 989
Db 1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1317
QY 990 CGAACACAGAACAAAGACCGCAGGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049
Db 1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1257
QY 1050 CCAGGGGGGCAAAAGGAGCGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1109
Db 1256 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1197
QY 1110 AAAAAGGGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1169
Db 1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1137
QY 1170 AAGCGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1229
Db 1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1077
QY 1230 CAAAGGAGAACCCCAAG 1247
Db 1076 RRRRRRRRRATCGCAG 1059
```

RESULT 2

```
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Iandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,999
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/710,561
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 25043-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 816/474-9050
; TELEFAX: 816/474-9057
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
```

```
; MOLECULE TYPE: DNA (genomic)
```

```
; HYPOTHETICAL: NO
```

```
; ANTI-SENSE: NO
```

```
; ORIGINAL SOURCE:
```

```
; ORGANISM: Staphylococcus aureus
```

```
; STRAIN: UT0007
```

```
; US-08-931-999-4
```

```
Query Match 7.6%; Score 102; DB 3; Length 6755;
```

```
Best Local Similarity 45.3%; Pred. No. 9.1e-14;
```

```
Matches 568; Conservative 0; Mismatches 675; Indels 11; Gaps 5;
```

```
QY 52 ACACACAGACAGCAAAACCCAGGACAGCGGAGCGCAACAGGAGAGAGAGAGAGAGAG 111
```

```
Db 5413 ACAGGCGAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 5472
```

```
QY 112 AACGGGCGCGCCGCAAGAGAGAGAACCCAGACCCCAAGAGAGGGGAGCGGCGACG 171
```

```
Db 5473 AAAGGCAAGCGAAGAGAGAGAGAAACAGAACCAAGCGACAGAGAGAAAAGGAA 5532
```

```
QY 172 GCAACCCAGCGCGGGGACACCAACCGGGGACAGGAGGCGCCCGGACACCGAGG 231
```

```
Db 5533 CCAACAGCAGGAGGAGGAGAAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 5592
```

```
QY 232 GGAGAACACCCCGACCAAGCCACCGGAAACAGCCGAGCCCGCGGACGAGAGAGCG 291
```

```
Db 5593 GGGGCAACACAGACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5652
```

```
QY 292 ACCAAAAGAAAGCAACAGAAACCGCCGACCCAGCGGAGCCAAACGAGAGAGAGAG 351
```

```
Db 5653 ACACAAAAAAGCAACAGCAACAA--AACCCAGAACACCGACAAAGAGAGAGAGAAA 5710
```

```
QY 352 CGCCCGAGAGCAGAGAGACGCACACACACAGAGAGGCGCCCGCGAGGAGAACCG 411
```

```
Db 5711 ACAACCAAGCAAGAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 5770
```

```
QY 412 AAACCGAGGCGAGCAACACCCGAGGAGGAGCCAGCCAGGAGAAACAGGAGAGAG 471
```

```
Db 5771 CCAGGACCCGACAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5830
```

```
QY 472 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
```

```
Db 5831 ACACCCAGAAACACCAACACACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 5890
```

```
QY 532 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 591
```

```
Db 5891 AGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5950
```

```
QY 592 GGAGAAAGAACCCGCGCAACCAACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
```

```
Db 5951 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6010
```

```
QY 652 AGACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
```

```
Db 6011 ACGAGAAAAACAAAGAAACAGGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 6070
```

```
QY 711 --ACCCACACAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
```

```
Db 6071 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6130
```

```
QY 769 GAAGCCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
```

```
Db 6131 GAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6190
```

```
QY 829 GAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
```

```
Db 6191 ACCAAAAACAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6249
```

```
QY 889 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
```

```
Db 6250 AAAAAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6309
```

```
QY 947 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
```


Db 467 ATGAAGGGGGGAGAAAAGGAAGAGAGAAAAGAGCGGAGAAAAGAGAAAAGAGAG 408
QY 948 GGAGGAAGCCAGACAGCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007
Db 407 GAAGGAAG 348
QY 1008 CGGAGGAAGCGGAG 1067
Db 347 AAAAGGGGGGAG 288
QY 1068 ACCGAG 1127
Db 287 AG 228
QY 1128 GAAGGGCCAAAGCACACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187
Db 227 AAAAGGAG 168
QY 1188 AAAGGAGCGCAG 1209
Db 167 TGGGGGGGGGAG 146

RESULT 5

US-09-270-767-1308
; Sequence 1308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1308
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1308

Query Match 6.1%; Score 82.6; DB 4; Length 1827;
Best Local Similarity 48.4%; Pred. No. 1.4e-09;
Matches 320; Conservative 0; Mismatches 334; Indels 7; Gaps 3;
QY 534 GGAAG 593
Db 1158 GGAAG 1217
QY 594 AGAAG 653
Db 1218 AAG 1277
QY 654 ACCGAG 713
Db 1278 AG 1277
QY 654 ACCGAG 713
Db 1278 AG 1337
QY 714 CAACCAAG 770
Db 1338 GAAAG 1397
QY 771 AGCCAG 830
Db 1398 GAAAG 1457
QY 831 AGAGGAG 890
Db 1458 AAG 1517
QY 891 GAG 950
Db 1518 AAG 1574

QY 951 GGAAGCCAG 1010
Db 1575 AGAGGGGAG 1633
QY 1011 CAG 1070
Db 1634 AAG 1693
QY 1071 GAAAG 1130
Db 1694 AGGAG 1753
QY 1131 GGGGAG 1190
Db 1754 GAAG 1813
QY 1191 G 1191
Db 1814 G 1814

RESULT 6

US-09-270-767-16590
; Sequence 16590, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16590
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16590

Query Match 6.1%; Score 82.6; DB 4; Length 1827;
Best Local Similarity 48.4%; Pred. No. 1.4e-09;
Matches 320; Conservative 0; Mismatches 334; Indels 7; Gaps 3;
QY 534 GGAAG 593
Db 1158 GGAAG 1217
QY 594 AGAAG 653
Db 1218 AAG 1277
QY 654 ACCGAG 713
Db 1278 AG 1337
QY 714 CAACCAAG 770
Db 1338 GAAAG 1397
QY 771 AGCCAG 830
Db 1398 GAAAG 1457
QY 831 AGAGGAG 890
Db 1458 AAG 1517
QY 891 GAG 950
Db 1518 AAG 1574
QY 951 GGAAGCCAG 1010

[illegible]

Search completed: October 21, 2004, 07:12:04
Job time : 143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2004, 05:53:10 ; Search time 1687 Seconds
(without alignments)
4082.128 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599

Perfect score: 1344

Sequence: 1 aagacacagcagagacgcg.....gagaagagagaagacccgc 1344

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
C 1	94.8	7.1	42999	9	US-09-799-462A-17
C 2	94.8	7.1	42999	10	US-09-836-911A-17
C 3	94.8	7.1	42999	10	US-09-738-630-73
C 4	94.8	7.1	42999	13	US-10-125-767-17
C 5	94.8	7.1	42999	14	US-10-151-081-17
C 6	94.8	7.1	42999	15	US-10-287-313-17
C 7	94.8	7.1	42999	15	US-10-694-17
C 8	94.8	7.1	42999	17	US-10-808-689-17
C 9	88.8	6.6	2318	10	US-09-814-353-21388
C 10	88.4	6.6	1468	15	US-10-369-493-39710
C 11	87.6	6.5	1840	15	US-10-017-161-2359
C 12	87.6	6.5	1840	15	US-10-292-798-2003
C 13	83.2	6.2	42533	15	US-10-004-113-43

C 14	81.4	6.1	6668	15	US-10-311-455-1670
C 15	80.6	6.0	1343	13	US-10-001-843-70
C 16	80.6	6.0	3259	13	US-10-001-843-71
C 17	79.8	5.9	1880	14	US-10-198-846-12701
C 18	79.2	5.9	1547	13	US-10-369-493-39721
C 19	78.2	5.8	792	13	US-10-027-632-168583
C 20	78.2	5.8	792	13	US-10-027-632-168584
C 21	78.2	5.8	792	13	US-10-027-632-168585
C 22	78.2	5.8	792	13	US-10-027-632-168586
C 23	78.2	5.8	792	13	US-10-027-632-168587
C 24	78.2	5.8	792	15	US-10-027-632-168588
C 25	78.2	5.8	792	15	US-10-027-632-168589
C 26	78.2	5.8	792	15	US-10-027-632-168590
C 27	77.2	5.7	1041	10	US-09-814-353-14764
C 28	77	5.7	868	17	US-10-437-963-42302
C 29	77	5.7	876	13	US-10-001-857-17
C 30	77	5.7	1186	15	US-10-017-161-2113
C 31	77	5.7	1186	15	US-10-292-798-1759
C 32	77	5.7	1225	15	US-10-017-161-2239
C 33	77	5.7	1225	15	US-10-292-798-1945
C 34	76.2	5.7	5258	14	US-10-198-846-10433
C 35	76	5.7	89625	13	US-10-087-192-2020
C 36	76	5.7	90614	18	US-10-331-053-64
C 37	75.4	5.6	1267	13	US-10-001-843-45
C 38	75.4	5.6	1297	15	US-10-017-161-2193
C 39	75.4	5.6	1297	15	US-10-292-798-1839
C 40	75.2	5.6	1706	15	US-10-369-493-39722
C 41	74.8	5.6	803	17	US-10-437-963-72176
C 42	74.8	5.6	1704	9	US-09-834-975-817
C 43	74.2	5.5	627	17	US-10-021-323-9336
C 44	74.2	5.5	1750	15	US-10-017-161-2107
C 45	74.2	5.5	1750	15	US-10-292-798-1753

ALIGNMENTS

RESULT 1

US-09-799-462A-17/c
; Sequence 17, Application US/09799462A
; Patent No. US20020160970A1

; GENERAL INFORMATION:

; APPLICANT: Hadlaczky, Gyula

; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF

; AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/799,462A

; FILING DATE: 10-Sep-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/835,682

; FILING DATE: 10-APR-1997

; APPLICATION NUMBER: 08/695,191

; FILING DATE: 07-AUG-1996

; APPLICATION NUMBER: 08/682,080

; FILING DATE: 15-JUL-1996

; APPLICATION NUMBER: 08/629,822

; FILING DATE: 10-APR-1996

; ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24601-402G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8403
TELEFAX: 858-587-5360
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 42999 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: <Unknown>
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-799-462A-17

Query Match 7.1%; Score 94.8; DB 9; Length 42999;

Best Local Similarity 45.3%; Pred. No. 2.1e-12;

Matches 568; Conservative 0; Mismatches 655; Indels 32; Gaps 5;

QY	90	ACAGAAAGAGAGAGGGGCGCCAAACCGGCGCCCGGACAAAGAGAGAGAACCCAGCCCAAGCAG	149
Db	25982	ATAGGAACAG	25923
QY	150	GGGACCGCGCAGCGGGCAGAGCAACCCAGCGCGCGGCGCACCCAAACCGGGGACAGGA	209
Db	25922	ACAGACAGAAAGACAGACAGACAGACAGAGAGAAACAGCGCAGAAAGAGAGAGAGAGAG	25863
QY	210	GGGCGCCGGAACCGAGAGAGAGAGAGAAACACCCCGGAACAGCCACCGGGAACAGCCGAG	269
Db	25862	AGAAACACAGACAGTAAAGGAG	25803
QY	270	CCGAG---CGCGGACAG	326
Db	25802	GACAGACACAGTGTAG	25743
QY	327	GACCAACGAG	386
Db	25742	GACAG	25683
QY	387	AGGCGCCCGGAG	446
Db	25682	AAGAGAGAAACAGGACAG	25623
QY	447	CAGGAGACAG	506
Db	25622	GAG	25563
QY	507	CCCCACGAGCGCCACACCCAG	566
Db	25562	ACACAGTGTAG	25503
QY	567	CAAGGAACCGGAG	626
Db	25502	ACGACAGAAAG	25443
QY	627	AAGCAGCCCAAG	686
Db	25442	CAGAGAAACAG	25383
QY	687	AAGAAAGACAGCGGACAG	746
Db	25382	CAGAAAGCGAG	25323
QY	747	AG	806
Db	25322	AAGACAG	25263

QY	807	GGAGAACACACAGAGAGCCCAAGGAGAGGACACACAGCGGCGAGCGGGAGCAGCCCGCC	866
Db	25262	AGACAGACAGAGAAAGGACACAGACAGAGAAAGACAGACAGAGAGAGAGAGAGAGAGAG	25213
QY	867	GCCACAGAAAGACAGAGAAAGACAGGAGAGACACAGAAAGAGGACCGCAGGGAGAGAGCC	926
Db	25212	-----AG	25159
QY	927	AGCAGAAAGGAGGACCGCGGAG	986
Db	25158	AGACAGACAGAGACAGACCGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	25099
QY	987	GGCGGAACACAGAAACAAAGAACCGCAGGACCGAGAGAGAGAGAGAGAGAGAGAGAGAG	1046
Db	25098	GGGAGAGACAG	25039
QY	1047	AACCCAGGGGGCAAAAGGAGCCGAAACCCAGCCACCCAGCCAGAGAGAGAGAGAGAGAG	1106
Db	25038	GAGAGAGACAGTGTAGAAACACACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	24979
QY	1107	CACAAAGGCGCAACAAAGGAGGAGGCGCAAAAGACACACCCCAACCAAGAGAGAGAGAG	1166
Db	24978	GNNAGAGAG-----AG	24926
QY	1167	GAAGAGCGCGAGAGAAAG	1223
Db	24925	GACAGACAGACAGGAG	24866
QY	1224	CCAGCCCAAG	1280
Db	24865	ACAGACAGAGAGAGACACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	24806
QY	1281	CAAGAAACAGAAAG	1335
Db	24805	AGAGAAAG	24751

RESULT 2

US-09-836-911A-17/c

; Sequence 17, Application US/09836911A

; Publication No. US20030033617A1

; GENERAL INFORMATION:

; APPLICANT: Hadlaczky, Gyula

; Szalay, Aladar

; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4350 La Jolla Village Drive, 6th Floor

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/836,911A

; FILING DATE: 17-Apr-2002

; CLASSIFICATION: <Unknown>

; PRIORITY INFORMATION:

; APPLICATION NUMBER: 08/835,682

; FILING DATE: 10-APR-1997

; APPLICATION NUMBER: 08/695,191

; FILING DATE: 07-AUG-1996

; APPLICATION NUMBER: 08/682,080

; FILING DATE: 15-JUL-1996

; APPLICATION NUMBER: 08/629,822

; FILING DATE: 10-APR-1996

; ATTORNEY/AGENT INFORMATION:

Db 25922 ACAGACAGAGAAAGACAGACAGACAGAGAGAAACAGCGAGAGAGAGAGAGAGAG 25863
Qy 210 GGGCCCGACACCGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269
Db 25862 AGAAACACAGACAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25803
Qy 270 CCCAG--CGCGGACAG 326
Db 25802 GACAGACACAGTGTAG 25743
Qy 327 GACCAACAG 386
Db 25742 GACAG 25683
Qy 387 AGGCCCCGAG 446
Db 25682 AAG 25623
Qy 447 CAGGAG 506
Db 25622 GAG 25563
Qy 507 CCCACAG 566
Db 25562 ACACAGTGTAG 25503
Qy 567 CAAGGAG 626
Db 25502 ACAG 25443
Qy 627 AAGCAG 686
Db 25442 CAG 25383
Qy 687 AAG 746
Db 25382 CAG 25323
Qy 747 AGAGGAG 806
Db 25322 AAGACAG 25263
Qy 807 GGAG 866
Db 25262 AGACAG 25213
Qy 867 GCACAG 926
Db 25212 -----AG 25159
Qy 927 AGCAG 986
Db 25158 AGACAG 25099
Qy 987 GGCGAG 1046
Db 25098 GGGAG 25039
Qy 1047 AACCCAGGGGGGAG 1106
Db 25038 GAG 24979
Qy 1107 CACAG 1166
Db 24978 GNNAGAGAG-----AG 24926
Qy 1167 GAAAGGGGGGAG 1223
Db 24925 GACAG 24866
Qy 1224 CCAAGCCAAAG 1280
Db 24865 ACAGACAG 24806

Qy 1281 CAAGAAACAGAAAGCCCGCGCAGCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1335
Db 24805 AGAGAAAG 24751

RESULT 5
US-10-151-081-17/c
; Sequence 17, Application US/10151081
; Publication No. US20030083293A1
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; Szalay, Aladar
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
; AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/151,081
; FILING DATE: 16-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/799,462
; FILING DATE: HEREMITH 05-MAR-2001
; APPLICATION NUMBER: 09/724,693
; FILING DATE: HEREMITH 28-NOV-2000
; APPLICATION NUMBER: 08/835,682
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: 08/695,191
; FILING DATE: 07-AUG-1996
; APPLICATION NUMBER: 08/682,080
; FILING DATE: 15-JUL-1996
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24601-402L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-450-8403
; TELEFAX: 858-587-5360
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-151-081-17

Query Match 7.1%; Score 94.8; DB 14; Length 42999;
Best Local Similarity 45.3%; Pred. No. 2.1e-12;
Matches 568; Conservative 0; Mismatches 655; Indels 32; Gaps 5;
Qy 90 ACAGGAG 149

QY	90	ACAGGAAGAGAGAGGGGCCAAAACGGGCCGCCGACAGAGAGNACCAGACCCCAAGCAG	149
Db	25982	ATAGGAACAGAGAGAGAGAGAGACAGAGAGACGACAGAAAAGAAAGAGAGAGGCGAG	25923
QY	150	GGGACCGCGCACCGCGCCAGGAGCAACCCAGCGCGGGGCACACCAACCGGGGACAGGA	209
Db	25922	ACAGCAGAGAAACAGAGACAGAGAGAAACAGGCAGAGAGAGAGAGAGAGAGAGAG	25863
QY	210	GGGCCCGGACAAACCGAGAGAGAGAGAAACAACCCCGGACAAGCCACCGGGAAACAGCCGGAG	269
Db	25862	AGAAACAGACAGTAAAGGAGAGAGAGAGAGACAGAGAGAGAAAGAGAAATAGACAGAA	25803
QY	270	CCGAG---CGCGCGAGCAGGAGCGACCAAAGAAAGCAACAGAAACCGCCCGACCGAGC	326
Db	25802	GACAGACACAGTGTAGAGAGGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	25743
QY	327	GACCAACGAGAGCAGAACCCCGAGCGCCCGGAGAGCAGAGAGAGCGACACACACACAG	386
Db	25742	GACAGAGAGAAAGACACAGACAGACAGAGAAAGACACAGACAGAGAGAGAGAGAGAG	25683
QY	387	AGGGCCCCGGAACGACGAACCAAGAACCGGGGCAGGACACCCGAGAGAGGACCCAG	446
Db	25682	AAGAGAGAAACAGGCAGAGAGAGAGAGCTAGCCAGAGAGAGAAACAGAAAGGTAGGGAGA	25623
QY	447	CAGGAGAACAGGAGAAAGCACCCCGAGCAGAGAGAGGACCGAGAGAGGACCCAGCAGGAAG	506
Db	25622	CAGAGAGAGACACACAGACAGACAGATGTGACAGCAGAGAGGAGAGTTAGACAGAGACAG	25563
QY	507	CCCCACGAGCGCCACACCCACGACCGCGAAGAGAGAAACCAAGAGGCCACAGGGCAGC	566
Db	25562	ACACAGTGAGAGAGACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	25503
QY	567	CAAGCAAGCCGAGAAAGGCCCGCGAGGAGAAAGAACCGCGCAACACCGGCGAGAGACAA	626
Db	25502	ACGACAGAAAGAGAGAGAGACAGACAGACAGAGAAACAGCAGAGACCGACAGAGAGAG	25443
QY	627	AAGCAGAGCCAAAAGAGCAGCAAAAGACCGGAAGACCGCGAGAGCCCAACAAAGCCCAAC	686
Db	25442	CAGAGAAACAGAGAGAAAGAGAGAAAGACAGAGAGCGAGAGAGGCGAGAGAGAGAGAA	25383
QY	687	AGAAAGACACCGGACGAGAGGAAACCCACCAACAGCGCCGACACGAGGCCACCGGCGCG	746
Db	25382	CAGAAAGCGGAGAGAGACAGAGAGAGACAGACAGATATAGCAGCGCAGAAAGAGAGT	25323
QY	747	AGAGGAGAACACCCACACCGCCCGAAGCCACAAAGACCAACCCCAACACAGAGAGAGAA	806
Db	25322	AGACAGAGATAGGCACAGAGAGAGAGACAGAGAGACACAGAAAGAGAAAGAGAGAGGC	25263
QY	807	GGAGAAACACCGAAGCCAAAGGAAGAGACACAGCGGCAGACGGGGAGCAGCCCGGCC	866
Db	25262	AGACACAGAGAAAGGGGACAGACAGAGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG	25213
QY	867	GCCACAGAGACAGAGAAAGCAGGAGAGACAGAGAAAGGGGACCGCAGGGAGAGAAAGGCC	926
Db	25212	-----AGAGAGAGAGAGAGAGAGAAACACACAGAAAGAGAGAGGACAGGGAGAGAG	25159
QY	927	AGCAGAAAGGAAGGACCGCGGAGGAAACCTAGAACCGCGAAAGAGAGAGCGCGAAGC	986
Db	25158	AGACAGACACAGACGACAGCGCAGAGAGAGAGAGGTAAAGCAAAAGATACACACAGAGA	25099
QY	987	GGCGACACAGAACCAAGAACCGCAGGAAACGGAAGAGAGGAGGAGAGAAACACAAAC	1046
Db	25098	GGGAGAGACAGAGAGACAGAGACAGACAGACAGAGAGAGAGAGTGAACACAGACAGACA	25039
QY	1047	AACCCAGGGGGGCAAAAGGGACCGGAAAAACCCAGCCCCACCGAGAGAGGAGACAGCCAG	1106
Db	25038	GAGAGAGACAGTGTAGAAACACAGCAAAAGAGAGGAGACAGAGAGAAACACACAGGGAGG	24979
QY	1107	CACAAAAGGCGACAAAGGGGAGGCCCAAGAGCACCCCAACAGAAACAGACAGGAGACA	1166
Db	24978	GNNAGAGAG-----AGAGAGAGAGAGAGACAGAAAGGGAGGGGAGACAGAGAGA	24926

Query Match	7.1%	Score	94.8	DB	17	Length	42999
Best Local Similarity	45.3%	Pred.	No. 2.1e-12				
Matches	568	Conservative	0	Mismatches	655	Indels	32
							Gaps
							5
90	ACAGGAAGAGAGAGGGGGCCAAAACGGGGCCGCCGACAGAGAGAGAACCCAGACCCCAACGAG	149					
25982	ATATGAAACAGAGACAGAGACAGAGACAGAGACGCGAATAAAGAAAGAGAGAGCGAG	25923					
150	GGGGACGGGGCACGGGGCCAGGAGCAACCCAGGCGGGGCACACCAACCGGGGACAGGA	209					
25922	ACAGACAGAGAAAGACAGACACAGACAGAGAGAAACAGGACAGAAAGAGAGAGAGAGAG	25863					
210	GGGCCCGGACAAACGGAGGAGGAGGAGAAACACCCCGGACAACCGCGGACACCCGGAG	269					
25862	AGAAACAGACAGTTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAGAGAAATAGACAGAA	25803					
270	CCGAG---CCCGGACGAGGAAGCGACCAAGAAAGACACAGAAACGGCCCGACCGAGCG	326					
25802	GACAGACACAGTGTAGAGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACAGACAG	25743					
327	GACCAACGAGACAGAGAACCCCGAGCGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	386					
25742	CTACTCTCAGAAACGACAGACAGACAGAGAAAGAGACAGACAGAGAAAGACAGAGAGAGAG	25683					


```
QY 891 GAGAGA-CCAGAAAAAGGGACCGCAGGGAGAAAGCCGAGCAGAAAGGAAAGGACCCGGG 949
Db 1928 AGAAGAGCGCAGAGAGGACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1987
QY 950 AGGAAAGCCAGAACCGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1009
Db 1988 AGAAAGAACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2047
QY 1010 GCAGGAACGGAGAGAGA-AGGGAGGAGAGAACACAAAAACACCCAGGGGGGCAAAAGGGA 1068
Db 2048 GAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2107
QY 1069 CCGAAAAACCCAGCCACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1125
Db 2108 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2167
QY 1126 GGGAGGGCCAAAAGCAGACCCACGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
Db 2168 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2227
QY 1186 AAAAGAGAGCGCAGAGAAAAAG 1208
Db 2228 AAGGAGGCACAAAGAGCAAGNG 2250

RESULT 10
US-10-369-493-39710
; Sequence 39710, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 39710
; LENGTH: 1468
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1468)
; OTHER INFORMATION: unsure at all n locations
US-10-369-493-39710

Query Match 6.5%; Score 88.4; DB 15; Length 1468;
Best Local Similarity 47.0%; Pred. No. 5.4e-11;
Matches 404; Conservative 0; Mismatches 447; Indels 8; Gaps 4;

QY 481 AGCAGCGAGAGGAGCAGCAGGAGAGGCGCCAGCAGGCGCCACACCCAGCCGCGGAAGAA 540
Db 210 AGAACAAAAACGCAACAGCGGATTCGCGCAAAACAGTTTTCGCAAGAAACCGGAGAGA 269
QY 541 GAAACCAAGAGGCCACAGGCGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 270 CACAAACACACACAGCGGAGAAAAACAGAGGGGGGGGAAACACACACAGAGAAAGAGA 329
QY 601 ACCGCGCAACAGCGGCAGAGACAAAAGCAGAGCCAAAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 330 GACACAGAAAAACACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 389
QY 661 ACCGCGCAGAGCCCAAGAGCCCAACCAAGAAAGACAGGAGAGAGAGAGAGAGAGAGAGAG 720
Db 390 CCAGACACAAAGAGACAAAGCAAAACACAGCGCCAAACACAGAAACACACACAGAGAGAG 449
```

```
QY 721 CAGSCCGAGACCGAGGCAACCGGGCCGCGAGAGAGAGAAACCCACACCCGCGGAGCCACAAAA 780
Db 450 GACAGACAGGGCCAGAGCAGAGAAACACAGAAATACCHACAGAGAGAGAGAGAGAGAGAG 509
QY 781 AGACCAACCCCAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 510 AACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
QY 841 AGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 570 AGCGCAACACACAAAGGAGACAAAACCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 629
QY 901 AARAAGGGACCGCAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 630 AAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 684
QY 961 ACAAGCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 685 AAAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 743
QY 1021 AGAAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 744 CGAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803
QY 1081 GCCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 804 GAAAGACCGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 862
QY 1141 CACACCCCAACGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 863 AGCGACACACACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 922
QY 1201 GAAAAAGGCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
Db 923 GACGCACAAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
QY 1260 CGAGGCCCAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
Db 983 ACACAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1042
QY 1320 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
Db 1043 AAAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1061

RESULT 11
US-10-017-161-2359/c
; Sequence 2359, Application US/10017161
; Publication No. US2003014368A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2359
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1840)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1640)
```



```
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (12)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (20)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (29)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (39)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (48)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (51)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (61)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (75)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (77)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (95)..(96)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (98)..(99)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (110)..(112)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (120)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (126)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (130)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (134)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (143)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
```

```
NAME/KEY: modified_base
LOCATION: (146)..(148)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (150)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (152)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (155)..(156)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (158)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (161)..(162)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (182)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (186)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (189)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (192)..(194)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (224)..(227)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (235)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (241)..(242)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (249)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (264)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (266)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (270)..(369)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (409)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
```

LOCATION: (429)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (437)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (481)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (489)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (528)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (578)..(579)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (605)

Query Match 6.5%; Score 87.6; DB 15; Length 1840;
Best Local Similarity 47.6%; Pred. No. 8.4e-11;
Matches 359; Conservative 0; Mismatches 379; Indels 16; Gaps 4;
QY 468 GCGAGCAGAGAGAGCGAGGAGGACCCAGCAGGAGGCGCCACAGCGCCACACCCCA 527
DB 1277 GAGAAGAAGGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1218
QY 528 CGACCCGGAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 587
DB 1217 NGGNGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1158
QY 588 GCGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647
DB 1157 AAAAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
QY 648 CAAAGACCGAGAGCGCCAGAGCCCAAGAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGG 707
DB 1097 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1038
QY 708 GAAACCC---AACCAACAGCCGAGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764
DB 1037 GAAAAAG 978
QY 765 CGCCGAGAGCCCAAGAGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 824
DB 977 GAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
QY 825 AAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 884
DB 917 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 858
QY 885 AGCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 944
DB 857 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 804
QY 945 CGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003
DB 803 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 744
QY 1004 AGAACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1063
DB 743 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 684

QY 1064 AGGACCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1123
DB 683 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 630
QY 1124 AGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1183
DB 629 AGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 570
QY 1184 AGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1217
DB 569 AAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 536

RESULT 12

US-10-292-798-2003/c
; Sequence 2003, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2003
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1840)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1640)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(1)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (12)..(12)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (20)..(20)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (29)..(29)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (39)..(39)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (48)..(48)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (51)..(51)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:

```

; NAME/KEY: modified_base
; LOCATION: (61)..(61)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (75)..(75)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (77)..(77)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (95)..(96)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (98)..(99)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (110)..(112)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (120)..(120)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (126)..(126)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (130)..(130)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (134)..(134)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (143)..(143)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (146)..(148)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (150)..(150)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (152)..(152)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (155)..(156)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158)..(158)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (161)..(162)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (182)..(182)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base

```

```

; LOCATION: (186)..(186)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (189)..(189)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (192)..(194)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (224)..(227)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (235)..(235)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (241)..(242)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (249)..(249)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)..(264)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (266)..(266)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (270)..(369)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (409)..(409)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (429)..(429)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (437)..(437)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (481)..(481)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (489)..(489)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (506)..(506)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (528)..(528)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (578)..(579)
; OTHER INFORMATION: a, t, c, g, unknown or other

```

Query Match 6.5%; Score 87.6; DB 15; Length 1840;
 Best Local Similarity 47.6%; Pred. No. 8.4e-11;

Sat Oct 23 17:44:56 2004

QY	943	ACGCGGAGGAAGCCAGACAAGCCGAAAGAGCAGCGGAAACGGCGCAACACAGAAACA	1002
Db	588	ACGGAGAGACATCGGCACATAGACAGAGGAGCCGCGGATACAAAGNAAAAGGAACA	647
QY	1003	-----AAGAACGGCAGGAACGGAAAGAGAGGGAGGAAGAACACAAACACCCAGGGGG	1057
Db	648	AACGGAAGATTGAGAAAATATGACGAGCGACGAAGCAACGACCGGAAACCCAGACCAGCGCG	707
QY	1058	GGCAAAAGGGACCGAAACCCAGCCACCCGAGAGGAGAGACAGCCAAACAAAAGGG	1117
Db	708	AGAGGCAGAGAAAGGAGCAGAGAACAAAAGCCGACAGAGAAAGCCAGACGAAAGAC	767
QY	1118	CAACAAAGGGGAAGGGCCAAAGCACACCCCAACGAAAGACAGGGACAGAAAAGCGCGC	1177
Db	768	AAGCACGAGCTACAGG--AGGAGCCAAAGATGAGAAAGAGAGAGAAAGAAACAC	825
QY	1178	CGAGAAAGAAAAGGAGCGGACAGAAAGGCCCCAGGGCGAGGGCCAGCCAAAGGAG	1237
Db	826	GAAGCAACAGACGACAGAACAGGAGAGAGAGAGAAACAGAGGGAGACGAAAGAGCAG	885
QY	1238	AACCCCAAGGACGACCGGACGAGGCCCAACCAACACCAAGCCCAAGAAACAGAAAAGCC	1297
Db	886	AGGAGAGAGAAACGAAAGTAGGGAGCCAGAGAAACGAAACGAGAGTACAAACAGNA	945
QY	1298	CGCGGCACGAAACAGACCCACACACAGCAGCAGAGAGAGAGAGA	1338
Db	946	CAGGGAAGAAAGAGACCACCAAGGACAAAAGNAGGAACACA	986

Search completed: October 21, 2004, 09:14:16
Job time : 1694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2004, 03:18:39 ; Search time 4597 Seconds
(without alignments)
10453.687 Million cell updates/sec

Title: US-10-019-341-4_COPY_256_1599

Perfect score: 1344

Sequence: 1 aagacacagcaggacacgc.....gagaagagagaagacccgc 1344

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*

2: gb_est2.*

3: gb_hc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gss1.*

9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120.2	8.9	2172	9	CL078016
2	119.6	8.9	1812	9	AG430199
3	118	8.8	997	9	CNS005TE
4	117.4	8.7	1723	9	AG430464
5	113.8	8.5	1389	9	AG386733
6	113.8	8.5	1511	9	AG350114
7	113.2	8.4	987	9	CNS00418
8	112.8	8.4	1407	9	AG346598
9	111.6	8.3	1397	9	AG346292
10	111.2	8.3	1424	9	AG346078
11	110.6	8.2	1542	4	BI252098
12	109.8	8.2	1542	4	AG346153
13	109	8.1	1318	9	AG324175
14	108.4	8.1	1319	9	AG324175
15	107.6	8.0	1923	9	AG360576
16	107.4	8.0	1260	9	AG278508
17	107.4	8.0	1582	9	AG346604
18	107.2	8.0	1418	2	BF683173
19	106.6	7.9	1402	9	AG324055
20	106.6	7.9	1519	9	AG279933
21	106.6	7.9	1595	9	AG365509
22	106.4	7.9	1189	4	BI230355
23	106.2	7.9	1362	9	CL082692
24	106.2	7.9	1795	9	CL043523

ALIGNMENTS

RESULT 1

CL078016

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Indels

Gaps

Gaps

C 25	106	7.9	1491	9	AG349727	Mus muscu
C 26	105.8	7.9	1518	9	CL078626	CH216-151
C 27	105.8	7.9	2129	9	AG393084	Mus muscu
C 28	105.6	7.9	1404	9	AG350266	Mus muscu
C 29	105.6	7.9	1439	9	CG743836	P036-1-H0
C 30	105.2	7.8	1571	9	CL066123	CH216-107
C 31	105	7.8	1673	9	CL082763	CH216-171
C 32	104.6	7.8	1515	9	AG346601	Mus muscu
C 33	104.6	7.8	1917	9	CL036131	Mus muscu
C 34	104.2	7.8	1459	8	BZ550911	pacsl-60
C 35	104	7.7	1360	9	AG448342	CH216-40P
C 36	104	7.7	1476	9	AG290071	Mus muscu
C 37	103.6	7.7	1786	9	CL082766	Mus muscu
C 38	103.6	7.7	2013	9	AG360864	CH216-171
C 39	103.4	7.7	1344	9	AG279247	Mus muscu
C 40	103.4	7.7	2142	9	CL082656	CH216-169
C 41	103.2	7.7	1133	8	CC205064	CH216-152
C 42	103.2	7.7	1463	9	CG757417	P052-3-H0
C 43	103.2	7.7	1464	9	AG430300	Mus muscu
C 44	103.2	7.7	2168	9	CG755980	P051-2-F0
C 45	102.8	7.6	1466	9	AG304877	Mus muscu

CL078016 2172 bp DNA linear GSS 31-DEC-2003
CH216-147G3_Sp6.1 CH216 Xenopus tropicalis genomic clone
CH216-147G3, genomic survey sequence.

CL078016
CL078016.1 GI:40533929

GSS.

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 2172)

Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert length: 175000 Std Error: 0.00

Seq primer: Sp6 ATTAGTGACACTATAG

Class: BAC ends

High quality sequence start: 436

High quality sequence stop: 531.

Location/Qualifiers

1. .2172

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/strain="Nigerian frog"

/db_xref="taxon:8364"

/clone="CH216-147G3"

/cell_line="Stock 248 F7A2, inbred N7"

/clone_lib="CH216"

/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis

BAC library"

Query Match 8.9%; Score 120.2; DB 9; Length 2172;

Best Local Similarity 46.0%; Pred. No. 3.8e-14;

Matches 483; Conservative 0; Mismatches 563; Indels 5; Gaps 2;

287 AGCGACCAAGAAAGCAAGACGCGCCGACGCGACGAGAGAGAGCC 346

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

[illegible]

RESULT 2
AG430199/c

LOCUS AG430199 1812 bp DNA linear GSS 03-JUN-2004
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-305C02.TJ, genomic survey sequence.
 ACCESSION AG430199
 VERSION AG430199.1 GI:48073262
 KEYWORDS GSS.
 SOURCE Mus musculus molossinus
 ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 BAC end Sequences of Library MSMg01
 2 (bases 1 to 1812)
 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 Direct Submission
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute, Bio Resource Center,
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp
 PRIMERS
 Sequencing : TJ
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.
 Location/Qualifiers
 1. . 1812
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-305C02.TJ"
 /sex="male"
 /tissue_type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"
 ORIGIN
 Query Match 8.9%; Score 119.6; DB 9; Length 1812;
 Best Local Similarity 46.3%; Pred. No. 5e-14;
 Matches 460; Conservative 0; Mismatches 528; Indels 5; Gaps 3;
 Qy 349 GAGGCCCGGAGACGAGAGACCCACACACACAGGAGGCCCGGAGCAGGAGAAC 408
 Db 1799 GAGAGAAAGNAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1740
 Qy 409 CAGAAACCGAGGGCAGAGCAACCCGAGGAGGAGCCAGCCAGGAGAACAGGAGAAC 468
 Db 1739 GAAGANAGGAGAGAGANAGAAAGAAAGAAAGNAGAGAGGAGAAAGAAAGGAGAG 1680
 Qy 469 CGGAGCAGAGAGAGAGCGGAGGAGGACCGAGAGAGGCCCCCAGAGGCCACACCCC 528
 Db 1679 AAGGAGGAAAGAGAGAGGAAAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAGA 1620
 Qy 529 GACCCGGNAGAAACCAAGAGGCCACAGGGCAGCCCAAGAGAGCCGAGAAAGGGCCG 588
 Db 1619 GGAAGAGGAGAAANANGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 Qy 589 CGAGGAGAAAGAACCCGCGCAACAAACGGGAGAGAGACAAAGAGAGCCCAAGAGAG 648
 Db 1559 AAAGNAGGAGAAAG 1500
 Qy 649 AAAAGACGAGAGACCGCAGAGCCCAAGAGCCCAAGAGAGAGAGAGAGAGAGAGAG 708

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	1137	AGAGAAGGAAAGAAAAGAAAAAGGGGAAGGAGNGCAAAAGAGAGAAAGAGAGAGAGAA	AG386733	1389 bp	DNA linear	GSS 03-JUN-2004	BAC end Sequences of Library MSMg01
Qy	998	GAACAAGAACCGCAGGAAACCGAAGAAGAGAGGAGGAGAGACACAAACCAACCAGGGG	AG386733/c				
Db	1077	GAATAAAGAAAGAAAGAAAG	AG386733.1	GL:47997938			
Qy	1058	GCATTAAGGAGCCGAAAGAACCCAGCCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG					
Db	1017	AATAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG					
Qy	1118	CAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG					
Db	957	GAATAAAG					
Qy	1178	CGAGAAAGAAAG					
Db	897	AAAAAAAAAAGAAAG					
Qy	1238	RACCCCAAAGGAACGGACCGGACGAGGCCCAACACCAAGCCCAAGAGAGAGAGAGAGAG					
Db	837	AAAAAAAAAANNANNAANAANAANAANAANAANAANAANAANAANAANAANAANAANA					
Qy	1298	CGCGCACCAACACAGACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG					
Db	777	AAAAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA					
RESULT 5							
LOCUS	AG386733/c						
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-201A07.TJ, genomic survey sequence.						
ACCESSION	AG386733						
VERSION	AG386733.1						
KEYWORDS	GSS.						
SOURCE	Mus musculus molossinus						
ORGANISM	Mus musculus molossinus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
REFERENCE	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y. BAC end Sequences of Library MSMg01						
TITLE	Unpublished						
JOURNAL	2 (bases 1 to 1389)						
REFERENCE	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y. Direct Submission						
TITLE	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)						
JOURNAL	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kunya Abe (abe@rtc.riken.jp) Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp						
COMMENT	PRIMERS Sequencing : TJ LIBRARY : pBACE3.6 Vector : EcoRI R.Site 1 : EcoRI R.Site 2 : EcoRI Location/Qualifiers 1..1723 /organism="Mus musculus molossinus" /mol_type="genomic DNA" /sub_species="molossinus" /db_xref="taxon:57486" /clone="MSMg01-305110.TJ" /sex="male" /tissue type="mixture of kidney and spleen" /clone_lib="MSMg01 Mouse Male BAC Library"						
FEATURES	source						
ORIGIN	Query Match 8.7%; Score 117.4; DB 9; Length 1723; Best Local Similarity 45.6%; Pred. No. 1.4e-13; Matches 402; Conservative 0; Mismatches 478; Indels 1; Gaps 1;						
Qy	458	GAGAAGCAGCGGAG					
Db	1616	GNAGAGAAAGAAAG					
Qy	518	CCCACACCCAGCAGCGGAG					
Db	1556	AG					
Qy	578	AGAAAGGCGCGGAG					
Db	1496	AG					
Qy	638	AAAGAGCAG					
Db	1436	GAAG					
Qy	698	GGGAG					
Db	1377	AAG					
Qy	758	CCCACACCGCGAG					
Db	1317	AAG					
Qy	818	GGAG					
Db	1257	AGAAAG					
Qy	878	CAG					
Db	1197	GAG					
Qy	938	AAGCAGCAG					

RESULT 6	
AG350114/c	
LOCUS	
DEFINITION	AG350114 Mus musculus molossinus DNA, clone:MSG01-146K01.TJ, genomic survey
ACCESSION	sequence.
VERSION	AG350114
KEYWORDS	AG350114.1 GI:47923424
SOURCE	GSS.
ORGANISM	Mus musculus molossinus Mus musculus molossinus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutherii; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

448	QY	AGGAGAACAGGAGAGCACC	CGCGGAGCAGAGAGGAGCGGAGAGGGAC	CAGCAGGAAGGC	507
1359	Db	AGAAGAAGGAGAGAGAAAA	GAGAGAAAAAAGGAGAAAGGANAAGAA	AAAGAGAGAGAA	1300
508	QY	CCACAGAGCGCCACAC	CCACAGCCCGGAGAGAGAAACCAAGAGGCC	CACAGGGCAGCC	567
1299	Db	AAGAAGNAGAAAAAGAA	AGAGAGAAAGAAAGAGAAAAAGAGAAAAAGAGAGAAAG		1240
568	QY	AAGGAAGCCGAGAAAGGG	CGCGAGGAGAGAAAGACCGCGCAACAA	CGGGCAGAGACAAAA	627
1239	Db	GAGAAAAGAGAGAAAAA	AGAGGAGAGAGAAAGAAAGAAAAAGAAAGAAAGAAAGAGAA		1180
628	QY	AGCAGAGCCAAAAAGAG	CAGCAAAAGCACCGCCAGAGCCCA	CAAGAGCCCAACCA	687
1179	Db	GNAGAGAAAAAGAACAAA	AGNAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA		1120
688	QY	AGAAAGACACGGGAC	CGAGAGGAAACCAACCAACAGCGCCGAGACCGGAGCCACCGGCCGA		747
1119	Db	AGAAAAGAAAAAAGAAAA	AAAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGGANGAGAGAGA		1060
748	QY	GAGGAGAAACACCCAC	ACCGCCGAGGCCCAAAAGAGACACCCCA	CAAGAGAGAGAGAG	807
1059	Db	AGAGAAAAAAGAGGAAA	AAAGAAAGAGAAAAAAGAAAAAAGAAAGAAAGGAGAGAAAAA		1000
808	QY	GAGAACACGGAAGCCAA	AGGAAGAGGACAAACAGCGGCAGACGGGAGCAGCCCGCCG		867
999	Db	AAGAGGAAAAAAGAAA	AGAAAGAGAGAGAGAAAAAGAAAGAGAGAGAAAGAAAGAAAGAA		940
868	QY	CCACAGAGACAGAAAA	AGCAGGAGAGACACAGAAAAAGGACCGCAGGGAGAAAGGCCA		927
939	Db	GAGNAAAAAGAGAGAAA	AGAGAAAAAAGAAAAAAGAAAAAAGAGAAAAAAGAAAAAGAAAA		880

```

928 GCAGAAAGGAAAGCGCCGCGAGGAAAGCCAGACAGCCGAAAGAGCGCGGAAACG 987
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
879 GAAGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
988 GCGGACACAGACAAAGACGCGAGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 1047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 AAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 762
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 ACCCAGGGGGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
761 -----AGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 713
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1108 ACAGAAAGGCGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
712 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1168 AAAAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 593
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1228 GCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
592 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1288 CAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
CNS00418/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC09C16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL066537
VERSION
1
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)

```

```

CNS00418 987 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC09C16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL066537
VERSION
1
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR09C16"
/clone_lib="RPCI-98"
/note="end : TET3"

```

```

FEATURES
source
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR09C16"
/clone_lib="RPCI-98"
/note="end : TET3"
ORIGIN

```

```

Query Match 8.4%; Score 113.2; DB 9; Length 987;
Best Local Similarity 27.5%; Pred. No. 9.5e-13;
Matches 147; Conservative 174; Mismatches 212; Indels 1; Gaps 1;

568 AAGGAAAGGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 627
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
966 ARAARGAARARARARARARARARARARARARARARARARARARARARARARA 907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
628 AGCAGAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
906 AARAAARARARARARARARARARARARARARARARARARARARARARARARA 847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
588 AGAAGACACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
846 GRGRRGRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRG 787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 GAGGAGAAACCCACACCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
786 ARGGGAAARARARARARARARARARARARARARARARARARARARARARARA 727
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 GAGAAACACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
726 AAAAARARARARARARARARARARARARARARARARARARARARARARARA 667
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 CCACAGAGACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
666 RRRRRGRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRG 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
928 GCAGAAAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 RRRARRAGAGARRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRG 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
988 GCGGAAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
547 GRRGGGGGGRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRG 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 ACCCAGGGGGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 AGAGARAGRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRG 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
AG346588/c
LOCUS
DEFINITION
Mus musculus molossinus DNA, clone:MSMg01-141M24.T7, genomic survey
sequence.
ACCESSION
AG346588
VERSION
AG346588.1
KEYWORDS
GSS.
SOURCE
Mus musculus molossinus
ORGANISM
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1407)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunuya Abe (abe@c.riken.jp).
Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@c.riken.jp

```


RESULT 10					
AG346078/c					
LOCUS	AG346078	1424 bp	DNA	linear	GSS 02-JUN-2004
DEFINITION	Mus musculus molossinus DNA, clone:MSG01-141B22.T7, genomic survey sequence.				
ACCESSION	AG346078				

Query Match	8.3%;	Score 111.2;	DB 9;	Length 1424;
Best Local Similarity	46.3%;	Pred. No. 2.5e-12;		
Matches 438;	Conservative 0;	Mismatches 500;	Indels 7;	Gaps 3;
398	GAAGCAGGACCAAGAAACACAGGGCGAGGACAAACCCGAAAGGAGGACCAAGGAGGAGAACAG	457		
1277	GAAGGAAGGAAAGAAGAAGGAAGGAAGAAGAAGAAGGAAGGAAGGAGAGAGAA	1218		
458	GAGAAGCACCGCGAGCAGAGAGAGAGACGAGAGGACCAAGCAGGAGAGGCCCCCAAGCGCG	517		
1217	AAAAAGGAAGAAGAAG	1158		
518	CCCACACCCACGACCCCGGAGAGAAAGAAACCAAGAGGCCCAAGGGCAGCCAGGAAGCCG	577		
1157	AAGA-AGGGAGAGGAG	1099		
578	AGAAAGGGCCGCGAGGAGAGAAAGAACCGCGCAACACCGGCGAGAGACACAAAAGCAGAGCCCA	637		
1098	GGAAAGAAAAAGAAAGAAAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	1039		
638	AAAGAGCAGCAAAAGACCGAAGACCCGAGAGCCCAAGAGCCCAACCAAGGAGAGAGAGAGAGAG	697		
1038	AAAAAGAGAA--AAAAAAG	981		
598	GGGACGAGAGGAAACCCCAACCAAGCGCCGAGAGCCGAGGACCCGGCGCGAGAGGAGAGACA	757		
980	AAAAAGAAAGAAAG	921		
758	CCCACACCGCGGAGCCCAAAAAAGACCAACCCCAACCAACAGAGAGAGAGAGAGAGAGAGAGAG	817		

Search completed: October 21, 2004, 07:09:42
Job time : 4606 secs

This Page Blank (uspto)